UNIVERSITÉ DE SHERBROOKE

TRF2: A new target for telomere dysregulation in human lymphoid cells

Par

Joannie St-Germain

Département de Microbiologie et d'Infectiologie

Mémoire présenté à la Faculté de Médecine et des Sciences de la Santé en vue de l'obtention du grade de Maître ès sciences (M. Sc.) en Microbiologie

Sherbrooke, Quebec, Canada
Date: 15 June 2018

Membres du jury d'évaluation
Pr. Raymund Wellinger, Département de Microbiologie et d'Infectiologie
Pr. Hans Knecht, Département de Médecine
Pr. Alfredo Menendez, Département de Microbiologie et d'Infectiologie
Pr. Martin Bisaillon, Département de Biochimie

RÉSUMÉ

TRF2: A new target for telomere dysregulation in human lymphoid cells

Par

Joannie St-Germain du Programme de Microbiologie

Mémoire présenté à la Faculté de médecine et des sciences de la santé en vue de l'obtention du diplôme de maîtrise ès sciences (M.Sc) en microbiologie, Faculté de médecine et sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, JIH 5N4

LMP1 est une protéine oncogène qui est observée dans les cellules tumorales dans le lymphome de Hodgkin associé au virus d'Epstein-Barr (EBV). LMP1 est responsable de l'interaction et / ou de l'interférence avec plusieurs voies de signalisation. Certaines de ces voies finiront par affecter l'activité transcriptionnelle de nombreux gènes en activant et en réprimant leurs promoteurs. Dans les cellules de lymphome hodgkinien EBV-négatives et EBV-positives, des expériences antérieures ont découverts des fusions-extrémitées chromosomiques, des translocations, des duplications chromosomiques et un dysfonctionnement des télomères. Le complexe Shelterin, composé des six sous-unités TRF1, TRF2, RAP1, TIN2, POT1 et TPP1, est associé à la formation de la structure de la boucle T-loop au niveau des télomères, formant ainsi un embout pour les chromosomes. Notre laboratoire a montré que l'expression de LMP1 induit une régulation négative des composants de Shelterin TRF1, TRF2 et POT1 au niveau transcriptionnel. Dans le complexe Shelterin, le TRF2 s'est avéré être la principale protéine affectée par l'induction de LMP1. Il a été montré que lorsque la protéine TRF2 conduite par myc est réintégrée dans une lignée cellulaire exprimant la protéine oncogène LMP1, cette LMP1 n'a plus d'effet sur sa régulation. Mon projet de maîtrise visait à comprendre comment LMP1 régule TRF2 au niveau de son promoteur. Notre hypothèse est que LMP1 interfère avec TRF2 au niveau du promoteur puisque les niveaux d'ARN de TRF2 sont réversibles lorsque LMP1 est supprimée. Plusieurs constructions de promoteurs de TRF2 fusionnées avec un gène rapporteur GFP ont été introduites dans une lignée cellulaire LMP1-tet-OFF et le niveau d'expression du gène rapporteur a été analysé par Western blot. Par cartographie appropriée de la région promotrice, un fragment de 30 pb du promoteur TRF2 a été identifié jouant un rôle central dans sa régulation. En utilisant la sequence du promoteur de 30 pb, une sonde radiomarquée a été créer. Un EMSA a été réalisé avec des extraits nucléaires provenant de cellules qui ont été induites ou non induites par LMP1. Les résultats révèlent des bandes distinctes et spécifiques pour des extraits nucléaires induits par LMP1 et non induits par LMP1 in vitro. Un séquençage de l'ARN avec un extrait cellulaire total a été réalisé pour obtenir des données sur les gènes qui sont régulés à la baisse et régulés à la hausse en présence de LMP1. Cela a été fait pour déterminer la composition dynamique cellulaire en l'absence et la présence de LMP1. Avec ces nouvelles données, une région spécifique sur le promoteur de TRF2 pour la régulation a été analysée par chromatographie d'affinité d'ADN et spectrométrie de masse. La découverte de protéines régulatrices du promoteur TRF2 aura un grand impact sur la compréhension des mécanismes régulateurs en jeu pour l'abondance de TRF2 in vivo.

Mots Clés: TRF2, LMP1, Promoteur, régulation négative

SUMMARY

LMP1 is an oncogenic protein that is observed in tumor cells of Epstein-Barr virus (EBV)-associated Hodgkin's lymphoma. LMP1 is responsible for interacting and/or interfering with multiple signaling pathways. At least some of those pathways will eventually affect transcriptional activity on many genes by activating and repressing their promoters. In EBV-negative as well as EBV-positive Hodgkin lymphoma cells, previous experiments uncovered chromosomal end-fusions, translocations, chromosomal duplications and telomere dysfunction. The Shelterin complex, consisting of the six subunits TRF1, TRF2, RAP1, TIN2, POT1 and TPP1, is associated with the formation of the T-loop structure at telomeres, thus forming an end cap for chromosomes. Our lab showed that expression of LMP1 induces a downregulation of the Shelterin components TRF1, TRF2 and POT1 at the transcriptional level. In the Shelterin complex, TRF2 was found to be the major protein affected by the induction of LMP1.

It was shown that when a myc-driven TRF2 is reintegrated into a cell line expressing LMP1 oncogenic protein, that LMP1 had no longer effect on its regulation. My master's project focused on understanding how LMP1 regulates the TRF2 at its promoter level. Our hypothesis is that LMP1 interferes with TRF2 at the promoter level since TRF2 RNA levels are reversible when LMP1 is suppressed. Several TRF2 promoter constructs fused with a GFP reporter gene were introduced into a LMP1-tet-OFF cell line and reporter gene expression level were monitored using Western blot. In that way, a LMP1-induced transcriptional shutdown could be documented. By appropriate mapping of the promoter region, a 30bp fragment of the TRF2 promoter was identified to play a pivotal role in this regulation. Using this 30bp as a dsDNA radio-labelled probe, EMSA was performed with nuclear extracts from cells that were induced or uninduced to express LMP1. The results reveal distinct and specific bands for LMP1-induced and LMP1-uninduced nuclear extracts *in vitro*. An RNA seq on the total cell extract was performed to obtain data of the genes that are downregulated and upregulated in the prescence of LMP1. This was done to determine the cellular dynamic composition in the absence and presence of LMP1-induction. These new data outline a specific region on the promoter of TRF2 for regulation that was analyzed by DNA affinity chromatography and mass spectrometry. The finding of regulator proteins of the TRF2 promoter will have great impact in understanding the regulatory mechanisms at play for TRF2 abundance *in vivo*.

Keywords: TRF2, LMP1, GFP, Promoter Regulation, Downregulation

List of Tables and Figures

Tables:

Figures:

Table of Contents

List of Symbols, Abbreviations, and Acronyms

°C: Degrees Celsius
µg: Microgram
µl: Microliter
µM: Micromolar
3': 3' Extremity of a nucleotide, free OH group on the 3' carbon of ribose ring
5': 5' Extremity of a nucleotide, free OH group or phosphate ester on 5' carbon of ribose ring
ALL-RW: all-around the world PCR
ATM: Ataxia telangiectasia mutated
ATR: Rad3-related
B: Base(s)
BB: Binding buffer
Bp: Base pair(s)
BJAB: Burkitt lymphoma cell line
B&W: Bind and wash buffer
BSA: Bovine Serum Albumin solution
CHAPS detergent: (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)
cHL: Classical Hodgkin lymphoma
CMV: Cytomegalovirus
ddPCR: Droplet digital polymerase chain reaction
DDR: DNA Damage Repair
DMSO: Dimethyl sulfoxide
DNMT: DNA methyltransferase
DNA: Deoxyribonucleic acid
dNTP: Deoxynucleoside triphosphate
DTT: Dithiothreitol
EBNA1: EBV nuclear antigen 1
EBV: Epstein-Barr virus
ECL: Electrochemiluminescence or electro generated chemiluminescence
EDTA: Ethylenediaminetetraacetic acid
EMSA: Electrophoretic mobility shift assay
FBS: Fetal Bovine Serum
GC: Germinal center
GFP: green fluorescent protein
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HL: Hodgkin Lymphoma
HRP: Horseradish peroxidase
HRS: Hodgkin and Reed-Sternberg cells
JAK: Janus kinase
JNK: C-Jun N-terminal kinase
Kb: Kilobases
kDa: KiloDalton
KCl: Potassium chloride
LB: Luria Broth medium
LMP1: Latent membrane protein 1

LMP2A: Latent membrane protein 2 a
M: Molar
MCS: Multi-cloning site
Mg: Milligram
$MgCl_2$: Magnesium chloride
Min: Minute
Mini-prep: Mini-preparation of plasmid DNA
Ml: milliliter
Mm: millimeter
Mmole: millimole
MMS: Methyl methanesulfonate
mRNA: Messenger ribonucleic acid
MS: Mass Spectrometry
m/v: Mass per volume
NaCl: Sodium chloride
NaOAc: Sodium acetate
NE: Nuclear extract
NEB: New England Biolabs
NF-κB: Transcription factor nuclear factor κB
ng: Nanograms
NHL: Non-Hodgkin lymphoma
NHEJ: Non-homologous end joining
Nm: Nanometer
nM: Nanomolar
Nt: Nucleotide
ORF: Open Reading Frame
PBS: Phosphate-buffered saline
PBS-T: Phosphate-buffered saline with 0.1% Tween-20
PCR: Polymerase Chain Reaction
pH: Measure of acidity or basicity of a solution
PMSF: Phenylmethylsulfonyl fluoride
PNK: Polynucleotide 5'-hydroxyl-kinase
POT1: Protection of telomeres 1
RAP1: Repressor/activator protein 1
RNA: Ribonucleic acid
RNase: Ribonuclease
RPM: Revolutions per minute
RPMI 1640: Roswell Park Memorial Institute medium
RS: Reed-Sternberg cells
RT: Room Temperature
RT-qPCR: Real-time PCR quantification
STAT: Signal transducer of activation
SDS: Sodium dodecyl sulfate
SDS-PAGE: SDS-Polyacrylamide Gel Electrophoresis
TAE: 40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA pH 8.0
TBE: Tris base, boric acid and EDTA

TCA: Trichloroacetic acid
TERC: Telomerase RNA Component
TERT: Telomerase Reverse Transcriptase
Tet: Tetracycline
TIN2: TRF1-interacting nuclear factor 2
TPP1: Tripeptidyl-peptidase I
TRE: Transactivator response element
TRF1: Telomeric repeat-binding factor 1
TRF2: Telomeric repeat-binding factor 2
tTa: Tetracycline transactivator
v/v: Volume per volume
V: Volt
VP16: Herpes simplex virus protein vmw65

Acknowledgments

I'd like to thank Raymund Wellinger for accepting me in his lab for giving me the opportunity to work with such a great mind in the field of telomere biology. Not only was he an exceptional mentor but he also helped me with my greatest fear, public speaking. I would never have imagined that I would be able to so confidently present my project in both my master seminars and speak at the 4th Canadian Symposium on Telomeres and Genome Integrity Symposium. I would also like to thank Bruno Lemieux for supervising my project and also giving me great expert advice. I'd like to especially thank Erin for being a great office partner and also correcting my memoire. I'd like to thank Dr. Hans Knecht for being my co-director in this project. I would also like to thank Pr. Alfredo Menendez and Pr. Martin Bisaillon for agreeing to be part of my jury.

Introduction

The goal of this study is to understand the impact of LMP1 on telomeres, more specifically, the telomeric protein TRF2, in the context of Hodgkin Lymphoma (HL). The study will focus on TRF2's promoter, and the regulation thereof in the context of a virus-associated cancer.

Importance of Telomere Function and the Shelterin Complex

Telomeres are unique structures found at the physical ends of linear eukaryotic chromosomes, which have multiple functions in preserving chromosomal stability, including protecting the ends of chromosomes from degradation and end to end fusion (Bailey & Murname, 2006). Telomeric functions depend on three important factors: telomeric DNA, the shelterin complex and telomerase. In humans and mice, the telomeric ends are composed of 2-30 kb of double-stranded TTAGGG repeats with the termini carrying a long 75-200 nt single-stranded overhang (Figure 1). The telomeric DNA is protected by the shelterin complex, which is comprised of several specific proteins that either bind telomeric DNA directly or are associated with telomeric chromatin (Maciejowski & de Lange, 2017). More specifically, the shelterin complex consists of the proteins TRF1, TRF2, POT1, TINT1/PTOP/PIPI (TPPI), RAP1, and TIN2 (Figure 1) (Palm & de Lange, 2008).

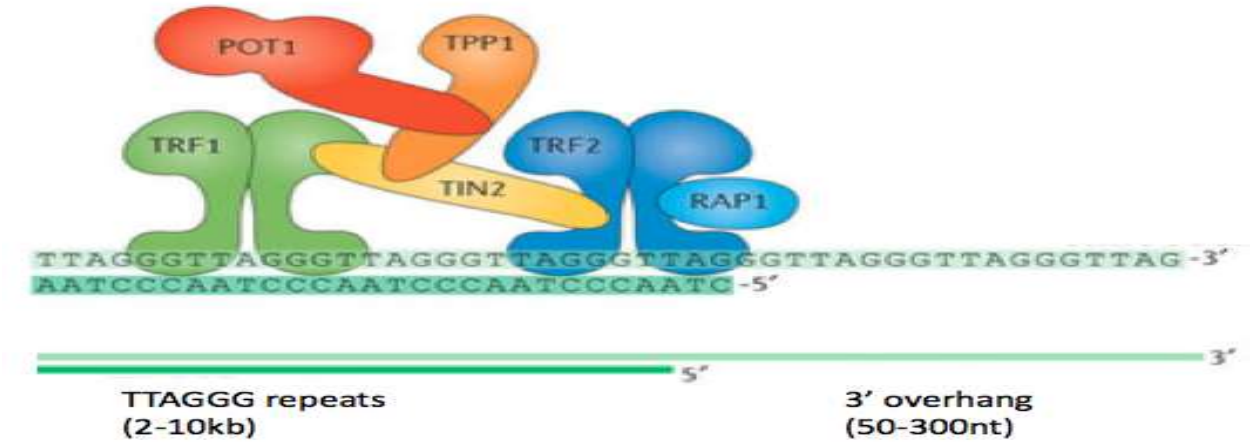


Figure 1: Telomeric structure.
A protein complex called shelterin forms the constitutive telomere architecture that is required for vital telomere function. The shelterin complex consists of six proteins, TRF1, TRF2, TIN2,

Rap1, TPP1 and POT1. TRF1 and TRF2 directly bind double strand telomeric DNA with high affinity, TIN2 binds both TRF1 and TRF2 to stabilize their association; Rap1 interacts with TRF2 and improves its selective binding to telomeric DNA while POT1 binds single strand telomere DNA. The 3' single strand overhang facilitates T-loop formation (Adapted from Maciejowski & de Lange, 2017).

The shelterin complex is recruited to the telomeres through TRF1 and TRF2 binding to the double-strand telomeric DNA and TIN2 that serves as a linker between the two homodimers. POT1 binds to the single stranded DNA and is linked to TRF1 and TRF2 through TPP1, which is associated with TIN2 (Maciejowski & de Lange, 2017).

The six components of the shelterin complex are responsible for formation of the t-loop structure (Griffith et al., 1999) which protects telomeres by repressing the DNA damage response (DDR) and repair pathways: the ATM kinase signaling pathway and non-homologous end joining (NHEJ) (Benarroch-Popivker et al., 2016). T-loops are formed through strand invasion of the single-stranded (ss) 3' telomeric overhang into the double-stranded (ds) part of the telomere, thereby forming a three-way junction at the base of the t-loop (Schmutz, Timashev, Patel & de Lange, 2017). This way, a telomeric DNA end is shielded from the DDR factors. The 3' overhang required for the t-loop formation is made by exonucleolytic degradation of the 5' ends of the telomeres, which therefore, shortens telomeres by about 50 bp per cell division (de Lange, 2005). Telomere shortening can be counteracted by telomerase which adds telomeric repeats at the end of the chromosome (Greider & Blackburn, 1987). Telomerase is a reverse transcriptase that synthesizes telomeric DNA using the integral RNA as the template and the 3'end of the telomere as the primer (Chan & Blackburn, 2002). Human telomerase is composed of a catalytic subunit TERT, an RNA component TERC and multiple accessory factors required for the activity of telomerase and its recruitment to telomeres (Cohen et al., 2007). During human development, telomerase will become deactivated through the downregulation of TERT (Holysz, Lipinska, Paszel-Jaworska & Rubis, 2013). As a result; somatic cells will undergo telomere shortening after each cellular division (Maciejowski & de Lange, 2017). This process of telomeric shortening will lead to the loss of telomeric DNA, chromosomal deprotection and activation of DDR pathways, which in turn leads to cell cycle arrest, senescence or apoptosis (Oeseburg, de Boer, van Gilst & van der Harst, 2009). Unlike normal cells, cancer cells will be under constant stress such as oncogenic stress, genomic instability and activation of apoptosis pathways. However, these cells escape apoptosis via genetic inactivation of certain pro-apoptotic

genes such as inactivation of BH3-only or caspase genes (Boyd et al., 1995). As a result, many cancer cells accumulate chromosomal abnormalities that are the direct result of telomere dysfunction due to inhibition of apoptosis (Mai & Garini, 2005).

TRF2 is the main shelterin component required for proper configuration of the telomeric t-loop, repression of DNA damage repair mechanisms at telomeres, and hence, protection of telomeres from fusions during cellular division. Deletion of the TRF2 protein has been shown to lead to chromosomal end fusion, DNA bridges, lagging chromosomes and giant, dicentric chromosomes during metaphase and anaphase (Figure 2) (van Steensel, Smogorzewska & de Lange, 1998; de Lange, 2005). This shows the importance that the shelterin complex plays in telomeric integrity and genome stability overall.

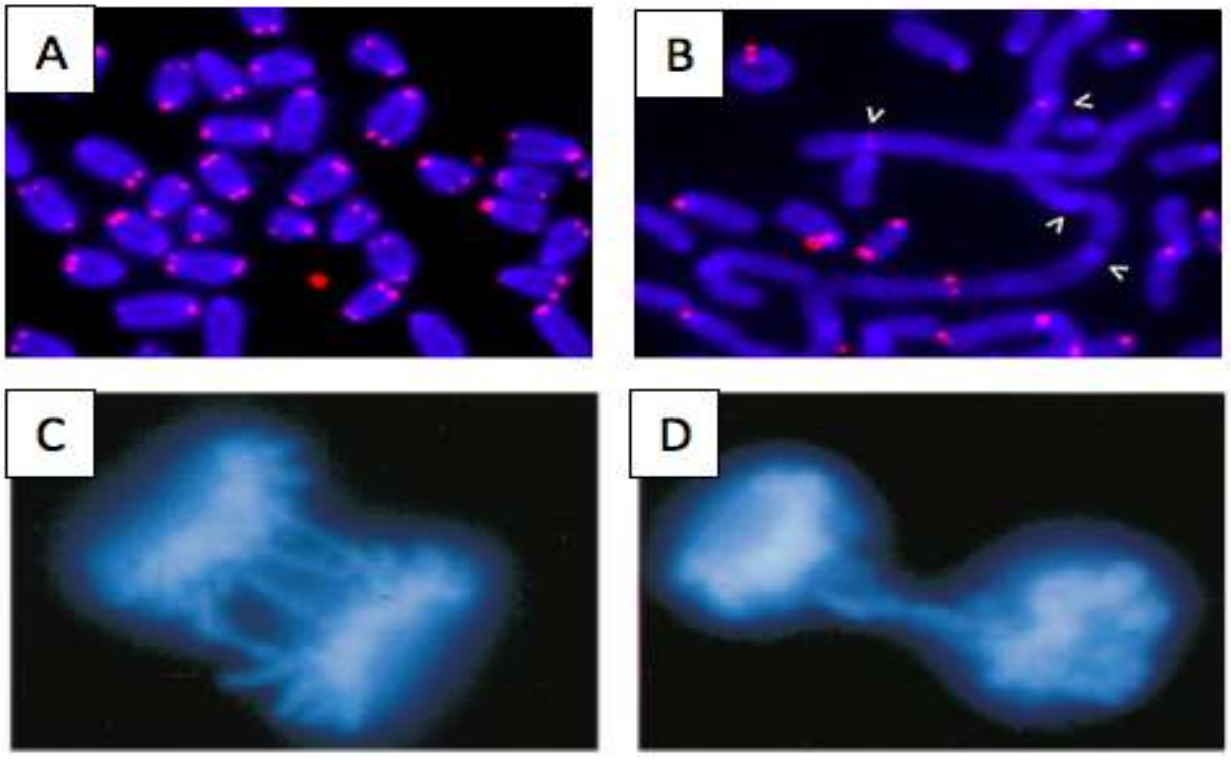


Figure 2: Cellular effects of TRF2 mutations.
a) Metaphase chromosomal spread, b) Metaphase chromosomal spread with removal of TRF2 by shRNA displaying DNA fusion, c) and d) Anaphase cells displaying TRF2 mutant missing its Myb domain induced bridges and lagging chromosomal in which the DNA was stained with DAPI (Adapted from van Steensel et al., 1998, Rai, Chen, Lei & Chang, 2016)

Telomere Integrity and Cancer

Due to the lack of telomerase activity in somatic cells, telomeres will shorten with each cellular division (Watson, 1972). This is a consequence of the telomeric end replication problem. In eukaryotes, DNA replication occurs via the leading and lagging strand synthesis. The leading strand is replicated in a continuous fashion by extension of a single RNA primer. The opposite lagging strand is replicated in short sections called Okazaki fragments and therefore requires multiple RNA primers. (Badaracco, Bianchi, Valsasnini, Magn & Plevani, 1985; Frick & Richardson, 2001). These primers are subsequently substituted with DNA due to the extension of the downstream Okazaki fragment (Balakrishman & Bambara, 2013). When the terminal RNA primer at the telomeric 3' end is removed, it cannot be replaced by DNA resulting in a short 3'-overhang. This process will be repeated in every consecutive round of replication, leading to the decrease in telomere length. Eventually, telomeres will become too short for shelterin proteins to bind, causing a loss of telomere protection and activation of DDR signaling pathway (Harley, Futcher & Greider, 1990). The DDR will in turn trigger activation of tumor suppressor mechanisms such as ATM, NHEJ, p53 and p16, driving cells into G1 arrest and inducing senescence or apoptosis (Figure 3) (d'Adda di Fagagna et al., 2003). Therefore, these tumor suppressor mechanisms are in place to eliminate and prevent the growth of cancerous cells (Yao & Dai, 2014).

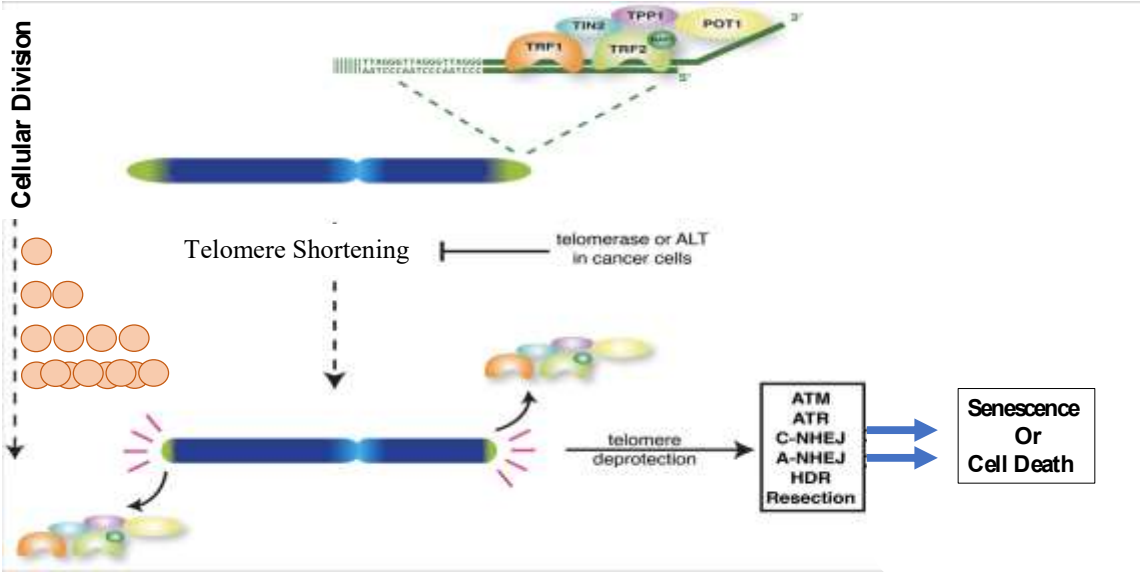


Figure 3: Telomeric integrity and role of the shelterin complex during cellular divisions. Telomeres are bound by the shelterin complex and with each cellular division, telomeres get shorter. If telomeres get too short, the shelterin complex will not bind, causing telomeric deprotection and activating the DDR pathway forcing the cells into senescence or apoptosis **(**Adapted from Jacobs, 2013).

There is a dynamic relationship between telomeres and genomic stability. The genome is susceptible to exogenous, endogenous and multiple other cellular processes that can compromise its integrity (Jackson & Bartek, 2009). Important mechanisms have been put into place to protect the genome from damage and instability (Nair, Shoaib & Sorensen, 2017). DDR mechanisms are the first responses to any genomic damage and tight regulation of these mechanisms is required (Ciccia & Elledge, 2010). Defects in DDR pathways such as a loss of function of p53, ataxia telangiectasia (ATM), or Rad3-related (ATR) kinases are associated with tumor initiation as well as tumor progression and metastasis (Broutas & Lieberman, 2014). Genomic and telomeric instability both interact closely in tumorigenesis. TRF2 mediates t-loop formation and maintenance, thus sequestering DDR pathways and protecting the telomeric ends from being processed by the NHEJ. When it comes to telomeric end deprotection due to short telomeres, the ATM-dependent pathway will become activated, inducing NHEJ and generating telomeric end fusions. However, functional DDR will drive cells into apoptosis, thus stopping propagation of genetically unstable cells. On the other hand, in the absence of the DDR pathway, cells will not

be forced into senescence or apoptosis but instead, will continue into division, initiating chromosomal instability (Figure 4) manifested by chromosomal translocations, duplications, chromosome loss, endomitosis and polyploidization (Davoli & de Lange, 2012). Pre-malignant cancer cells need to acquire a telomere maintenance mechanism during neoplastic transformation and for proliferation (Yakoob, Hu, Fan & Zhang, 1999). Telomerase reactivation may occur in late tumorigenesis as a result of ongoing genetic instability. It was recently found that a single nucleotide somatic mutation in the proximal promoter of telomerase reverse transcriptase (TERT) was enough for telomerase reactivation (Horn et al., 2013). This discovery emphasized the importance of promoter mutation in the progression of genomic instability in cancer development.

The cell depends on the ability of the DDR checkpoint to induce cell cycle arrest when critically short telomere length is reached. The effectiveness of DDR can either inhibit or promote cancer. Therefore, short telomeres are the earliest detectable hallmark of cancer. One form of cancer in which short telomeres are noticeable in its early stage is Hodgkin lymphoma (Jacobs, 2013).

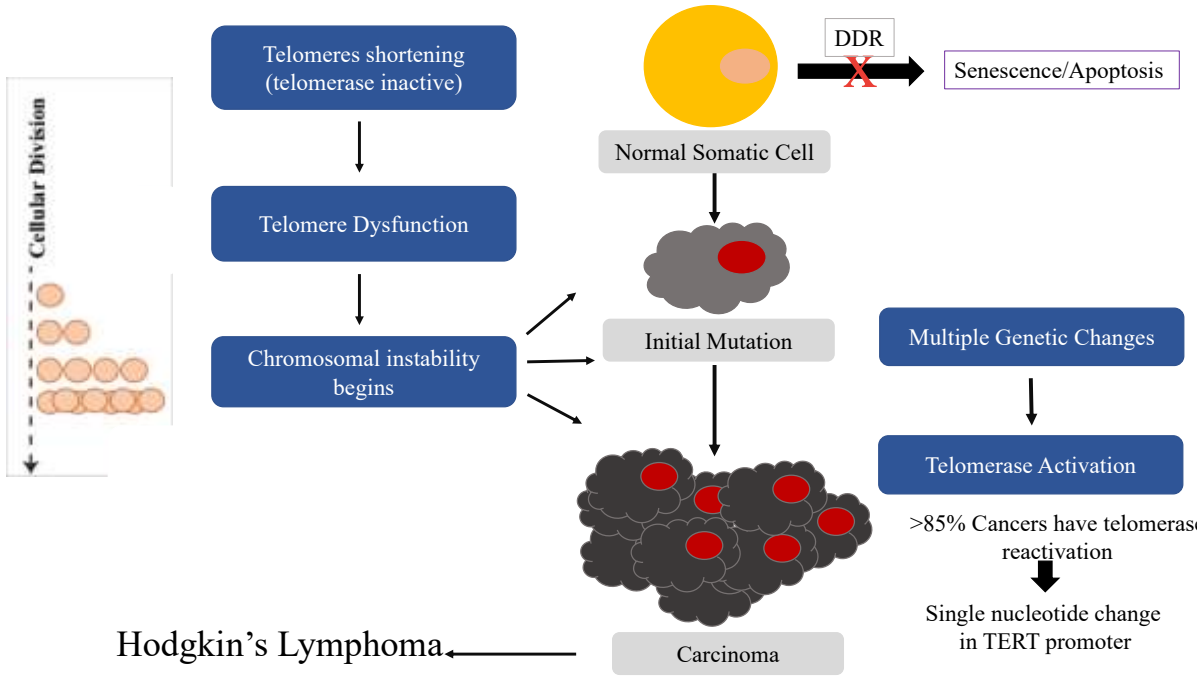


Figure 4: Model of cancer initiation.
Cancer development caused by mutation in the DDR pathway when short telomeres are present. Short telomeres will progress with each cellular division, causing chromosomal instability and genomic instability. Telomerase will be reactivated during late tumorigenesis.

Hodgkin Lymphoma (HL)

Lymphomas are a group of blood cancers that develop in the lymphatic system. Malignant cells accumulate in the lymphatic system, eventually causing enlargement of the lymph nodes (Ganti, Albershardt, Iritani & Ruddell, 2015). There are two types of lymphomas: the classic Hodgkin lymphomas (cHL) and the Non-Hodgkin-Lymphomas (NHL). The latter are divided into B-NHL and T-NHL (Swerdlow et al. 2016). The Hodgkin lymphomas are characterized by the mononuclear Hodgkin (H) cells and the diagnostic binucleated or multinucleated Reed-Sternberg (RS) cells. H and RS cells are often summarized as HRS cells. HRS cells do not express the majority of proteins proper to B-cell immunoblasts. A typical characteristic of both are Hodgkin mononucleated cells and multinucleated Reed-Sternberg cells, but contrary to non-Hodgkin lymphoma, cHL does not express the majority of proteins proper to B-cell immunoglobulins. Instead, they display different molecules not normally expressed by B cells (Gobbi, Ferreri, Ponzoni & Levis, 2013). Interestingly, in cHL several transcription factors are deregulated, but not in nHL. The mononucleated and multinucleated cells lack transcription factors such as OCT2, BOB.1 and PU.1, which are required for activation of the promoter or enhancers of immunoglobulin genes. These differences between the immunophenotypes and gene expression clearly puts Hodgkin lymphoma in a different category of lymphomas (Stein & Bob, 2009).

Cellular Origin of B cells including HRS cells

B cells are lymphocyte cells that play a crucial role in the immune system response. The main function of B lymphocytes is to produce specific antibodies. Their production and development begins by stimulation by exogenous factors such as bacteria or viruses that activate the production of hematopoietic stem cells in the bone marrow (Tobon, Izquierdo & Canas, 2013). Once exiting the marrow, all B cells are morphologically similar and also referred to as immature B cells (Gathings, Lawton & Cooper, 1977). Immunoglobulins play a critical part in the immune system by recognizing and binding to particular antigens. Antibodies are a form of immunoglobulin secreted by B cells to bind, neutralize, and send a signal to activate other defenses like macrophages to destroy foreign invaders (LeBien & Tedder, 2008). These immature B cells only become activated once the immunoglobulin binds to an antigen. This antigen binding induces B cell-activation, proliferation and differentiation in secondary

lymphoid tissues. These dynamic changes in gene expression and immunoglobulin mutation happen in the germinal center follicles of secondary lymphoid organs (Küppers, 2009). The B cells then undergo rapid proliferation to generate a large number of cells. Subsequently, immunoglobulins undergo somatic hypermutation to generate high-affinity antibodies unique to the antigen of the foreign invaders. This biological process also has a program that eliminates B cells not expressing the correct antibody and forces them into apoptosis (Figure 5). Only B cells harboring favorable mutations with specific antibodies will undergo class-switch recombination and differentiate to become plasma cells or memory B cells. A large number of genes are involved in the regulation of proliferation, differentiation and apoptosis, and can occasionally become dysregulated by errors during antibody gene modifications, thus causing genomic alterations. A series of checkpoints normally control the B-cells' selection for unfavorable mutations resulting in cellular apoptosis. In some cases, unfavorable B cells avoid apoptosis because of genetic modifications or due to the presence of the Epstein-Barr virus (EBV), thus resulting in oncogenic transformation, resulting in cancers such as Burkitt Lymphoma and Hodgkin Lymphoma (Klein & Dalla-Favera, 2008). In cHL the 5-year survival rate of patients that will become long-term survivors is 85%, but 15% of patients will die from a rapid reoccurrence of the disease or become resistant to drug-therapy (et al., 2011). HL is diagnosed by the presence of multinucleated Reed-Sternberg cells (Shanbhag & Ambinder 2018). With a better understanding of the processes involved in the formation of the Reed-Sternberg cells, we think new drug treatments for treating Hodgkin's lymphoma may emerge.

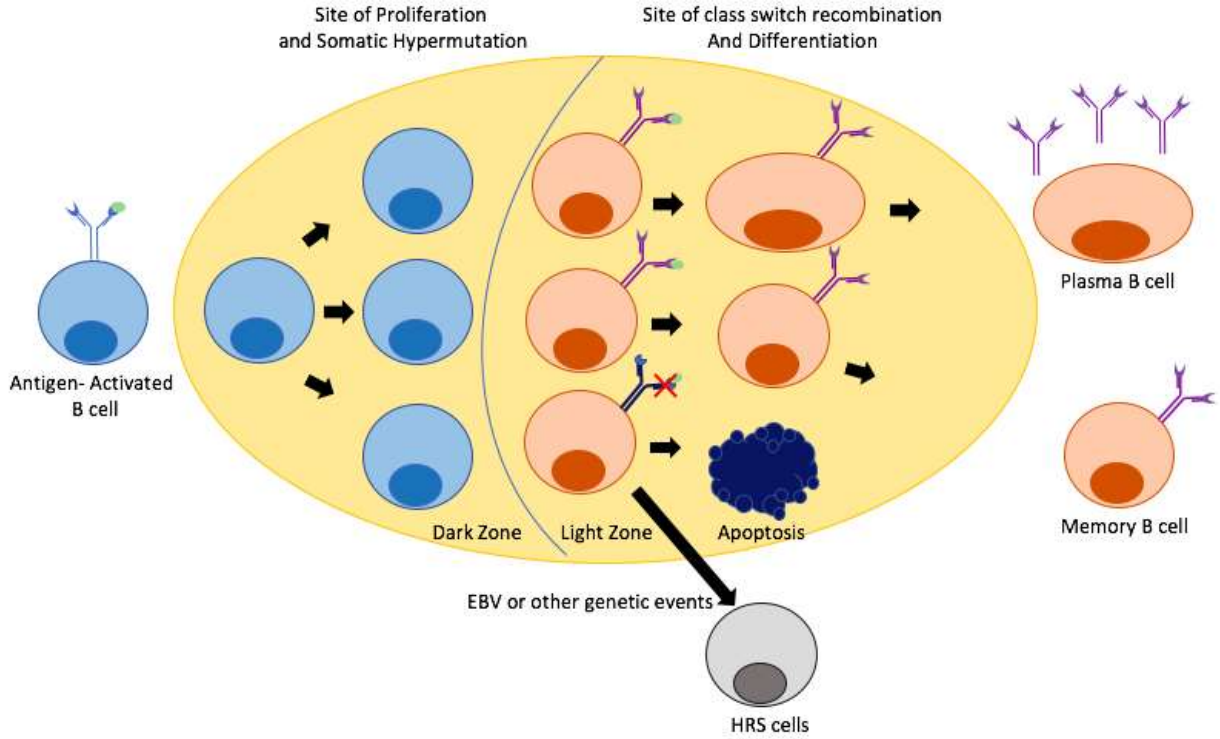


Figure 5: Origin of Hodgkin and Reed-Sternberg cells.
In the germinal center, the antigen-activated B-cells enter proliferation and somatic hypermutation. The somatic hypermutation will diversify B cell immunoglobulin receptors to adapt to the foreign invader, creating rearrangements in the immunoglobulin gene. The B-cells then go through antibody class switching recombination. This will allow the cells to have antigenic specificity with antibodies of different subtypes. The B-cells that have obtained a favorable rearrangement will continue to differentiate into plasma cells or memory B cells. Hodgkin and Reed-Sternberg cells (HRS) are cells that have acquired unfavorable rearrangements and should have undergone apoptosis, but because of genetic alterations or the presence of Epstein –Barr virus (EBV) derived proteins, they have survived (Adapted from Klein & Dalla-Favera, 2008).

Origin of Reed-Sternberg cells

HL is characterized by the presence of the mononuclear H cells, and bi or multinucleated Reed-Sternberg cells (HRS cells). It has been shown that HL cells have a large number of chromosomes with aberrations, genetic alterations, and chromosomal deletions (Knecht et al., 2008). A multitude of translocations have been identified in HRS cells and their nuclear architecture becomes progressively more disorganized as the number of nuclei increases. As well

as chromosomal disorganization, HRS also have an increased number of centrosomes, defective spindle formation, and genomic instability. The hallmark of HRS tumor cells is telomeric dysfunction (Knecht, Sawan, Lichtensztejn, Z., Lichtensztejn, D. & Mai, 2010). It is the main mechanism through which mononucleated Hodgkin cells become multinucleated Reed-Sternberg cells, in which the Hodgkin cells increase the disruption of their telomeres, genomic DNA and dynamically progress towards becoming Reed-Sternberg cells. These multinucleated cells are formed by endomitosis, a process in which the cells duplicate their chromosomes but this is not followed by cellular division. This phenomenon is associated with disturbed cytokinesis and jumping translocations due to severe telomere dysfunction. Hans Knecht et al. 2008, provided a 3D telomere analysis diagram on the transition of H to RS cells, providing a direct graphical demonstration for how an H cell becomes an RS cell (Figure 6). This 3D nuclear approach allowed the visualization of telomeric organization as the cells underwent transformation from an H to an RS cell. The mononucleated Hodgkin cells already have distinctive short telomeres, and telomere aggregates. As the cells undergo endomitosis, they will produce binucleated Reed-Sternberg cells, with more telomeric dysfunction such as telomeric aggregates, even shorter telomeres and some losses of telomeric signal (Knecht et al., 2008). This progression continues until the cell reaches the end stage multinucleated Reed-Sternberg cell, in which the cell no longer duplicates its genome, possibly due to extreme telomeric dysfunction and genomic instability (Knecht et al., 2010). Reed-Sternberg cells indeed show little proliferative capacity, thus demonstrating that Hodgkin cells could be the main cells with proliferative ability and the main reason for the generation of these Reed-Sternberg cells (Rengstl et al., 2013). Interestingly, in about 40-50% of cases of HL, patient biopsies of their H and RS cells have been found to express the Epstein Barr virus-encoded LMP1 oncoprotein (Knecht et al., 1993).

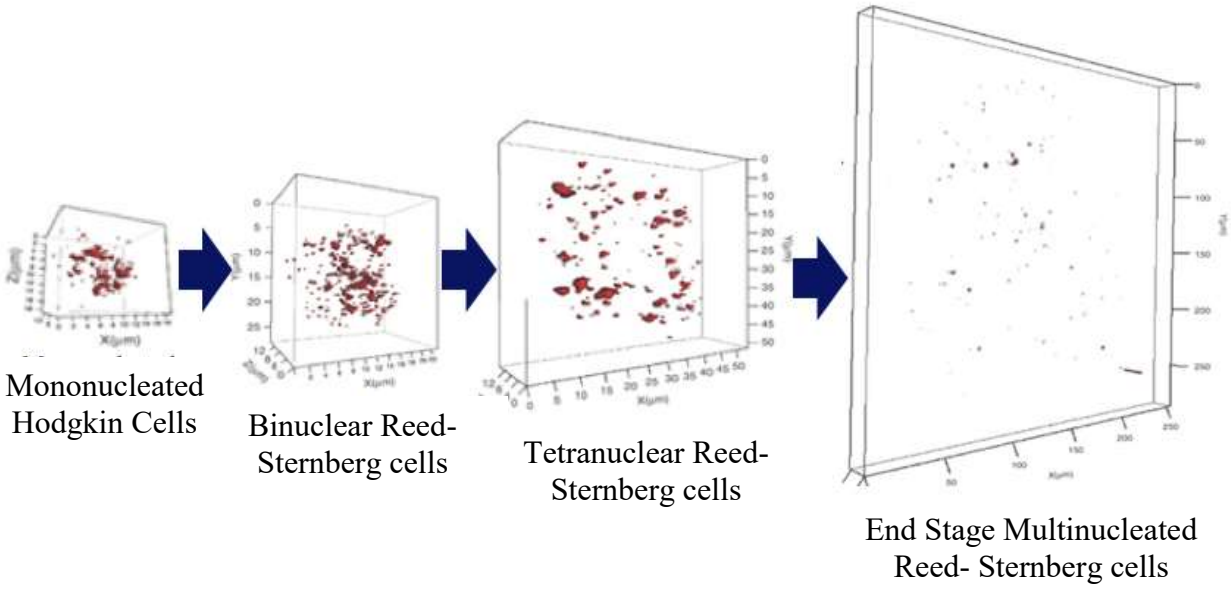


Figure 6: Correlation of telomere dysfunction and of RS-cell formation.
The model illustrates the evolution from a mononuclear H cell to a multinuclear RS cell. 3D reconstructed representative nuclei and their telomeres. A mononuclear H cell may either progress to a telomere-poor, binucleated RS cell or to a binucleated RS cell with a 'ghost' nucleus probably telomere-poor end-stage cells, or progress to a still telomere-rich, binucleated RS cell or large twisted H cell capable of further nuclear duplications. This may result in a trinuclear or tetranuclear RS cell. Depending on the 3D configuration of telomeres and the number of telomere aggregates, these RS cells may or may not be able to perform further nuclear division, resulting in multinucleated end-stage RS cells (Adapted from Knecht et al., 2008).

Pathogenic role of Epstein-Barr Virus and its trans-membrane protein called "Latent Membrane Protein 1" (LMP1)

Epstein-Barr virus (EBV) is a gamma-herpesvirus that establishes life-long sero-positivity and infects 90% or more of the world's population. Most people become infected at some point in their childhood, with a generally asymptomatic primary infection that can become symptomatic in their adolescence or adulthood, causing infectious mononucleosis (Hau & Tsao, 2017). Rarely, the disease incites a range of malignancies including nasopharyngeal carcinoma, lymphoproliferation disorders and Burkitt lymphoma (Kutok & Wang, 2006). In about 40% of cases of classical HL, the HRS cells are infected with the Epstein-Barr virus, thus indicating an importance of EBV in HL pathogenesis. EBV has a 170kb double-stranded DNA genome that forms an episome in the infected host cell, and exists as an extrachromosomal DNA wrapped

with histones. EBV will infect epithelial cells as well as B-lymphocytes (Riley et al., 2012). EBV does not usually replicate in B-lymphocytes, but instead establishes a latent infection, which is characterized by the limited expression of a subset of virus latent genes (Babcock, Hochberg & Thorley-Lawson, 2000). These subsets of individualized latent genes are the three main viral proteins called: EBNA1 (EBV nuclear antigen), LMP1 and LMP2A (Hau & Tsao, 2017). EBNA1 is a mandatory protein responsible for the replication of the viral genome and may have additional roles (Gahn & Sugden, 1995). LMP2A mimics an Ig receptor and recruits cytoplasmic kinases (Mancao & Hammerschmidt 2007). LMP1 is the main viral protein required for B cell immortalization and promotes the formation of Reed-Sternberg cells as shown by Knecht et al. 1999 (Figure 7).

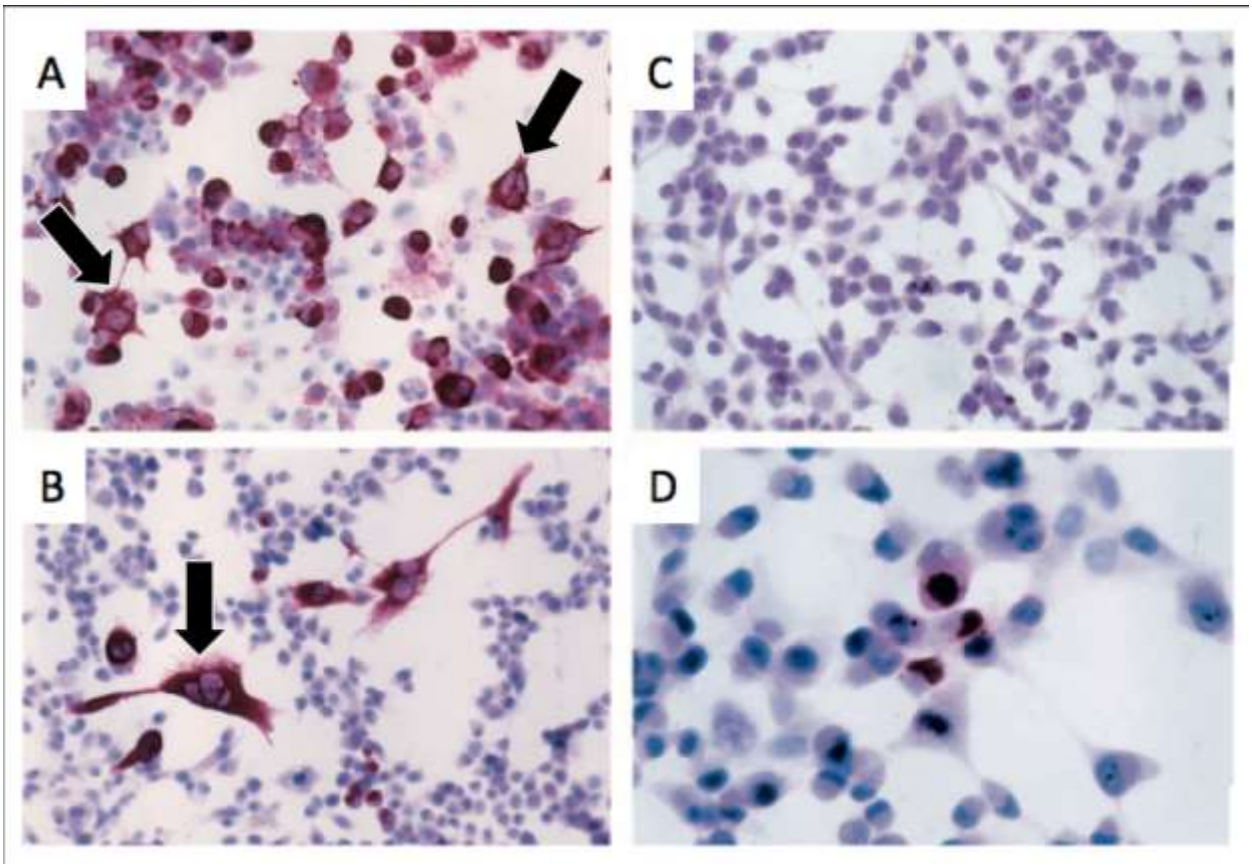


Figure 7: Histology of cells transfected with LMP1.
Impact of LMP1 expression in myelomonocytic HD-MyZ cells. A) represents cells that have been transfected with LMP1, showing evidence of multinucleated cells or cells larger than the normal surrounding cells. B) Immunostaining showing LMP1 expression C) Transfection cells with an empty vector D) Transfection with another viral protein EBNA1 with immunostaining revealing no change of cell morphology. Concluding that LMP1 is the viral protein of EBV causative of the multinucleated cells (Adapted from Knecht et al., 1999).

LMP1 is the major transforming protein of EBV and also plays a crucial role in the pleiotropic effects on cellular phenotypes and in the maintenance of EBV-induced tumours (Ahsan, Kanda, Nagashima & Takada, 2005). It is the first protein being produced during EBV latency establishment (Dawson, Tramountanis, Eliopolis & Young, 2002). Functionally, LMP1 resembles CD40, a member of the TNFR family, and can partially substitute for CD40 in vivo, providing both growth and differentiation responses in B cells (Brauninger et al., 2005). The LMP1 protein is a trans-membrane protein of 63 kDa and can be subdivided into three domains: (a) an N-terminal cytoplasmic tail (amino acids 1–23) which tethers and orientates the LMP1 protein to the plasma membrane; (b) six hydrophobic transmembrane loops that are involved in self-aggregation and oligomerization (amino acids 24–186); (c) a long C-terminal cytoplasmic region (amino acids 187–386) that possesses most of the molecule's signaling activity (Uchida et al., 1999). Two functional domains referred to as C-terminal activation regions 1 and 2 (CTAR1 and CTAR2) have been identified for their ability to activate the NF-κB transcription factor pathway (Eliopoulos, 1999). This effect contributes to the many phenotypic consequences of LMP1 expression including the induction of various apoptosis inhibition pathways and cytokine genes (Figure 8) (Huen, Henderson, Croom-Carter & Rowe, 1995). LMP1 contributes to 25% of transcriptional changes in HL cells transferase (Vockerodt et al., 2014). It will also modulate the expression of key tumor suppressor genes such as p53 and contributes to cell survival by activating several pathways such as NF-kappa B (NF-kB), JAK/STAT, and P13L/AKT pathways (Hoesel & Schmidt, 2013; Vaysberg, Lambert, Krams & Martinez, 2009). It is also involved in the loss of expression in mitotic spindle formation, disorganizes interphase microtubules, promotes chromosome breaks, promotes cells survival and B cell immortalization and is involved in promoter regulation through methylation by DNA methyl transferase (Vockerodt, Cader, Shannon-Lowe & Murray, 2014). Therefore, LMP1 contributes to cellular dysregulation and tumorigenesis.

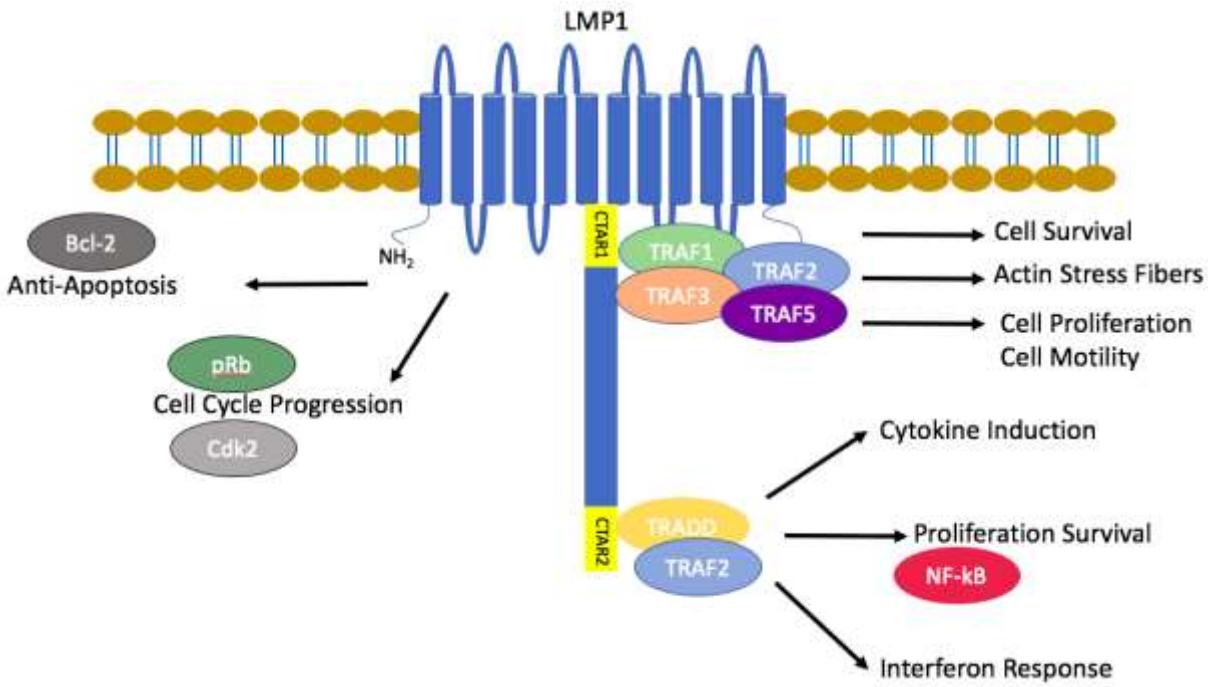


Figure 8: LMP1 structure and targeted pathway.
Expression in EBV-positive cases, contributing to NF-kB activation as it mimics an activated CD40 receptor which supports survival of the deregulated B lymphoma cells (Adapted from Brauninger et al., 2005).

Epstein-Barr virus positive Hodgkin Lymphoma

The EBV-negative and EBV-positive HRS cells show a distinct somatic mutation pattern underlining that the EBV infection happened in a B-cell in the germinal center (Thorley-Lawson & Gross, 2004). The EBV-positive HRS cells show a similar phenotype: B cell nuclear division becomes impossible with telomeric loss, telomeric shortening, end to end fusions and telomere-poor or telomere-free 'ghost' nuclei (Knecht et al., 2010). LMP1 expression was shown to be associated with multinuclearity in which many of the cells resembled RS cells (Figure 9) (Lajoie et al., 2015).

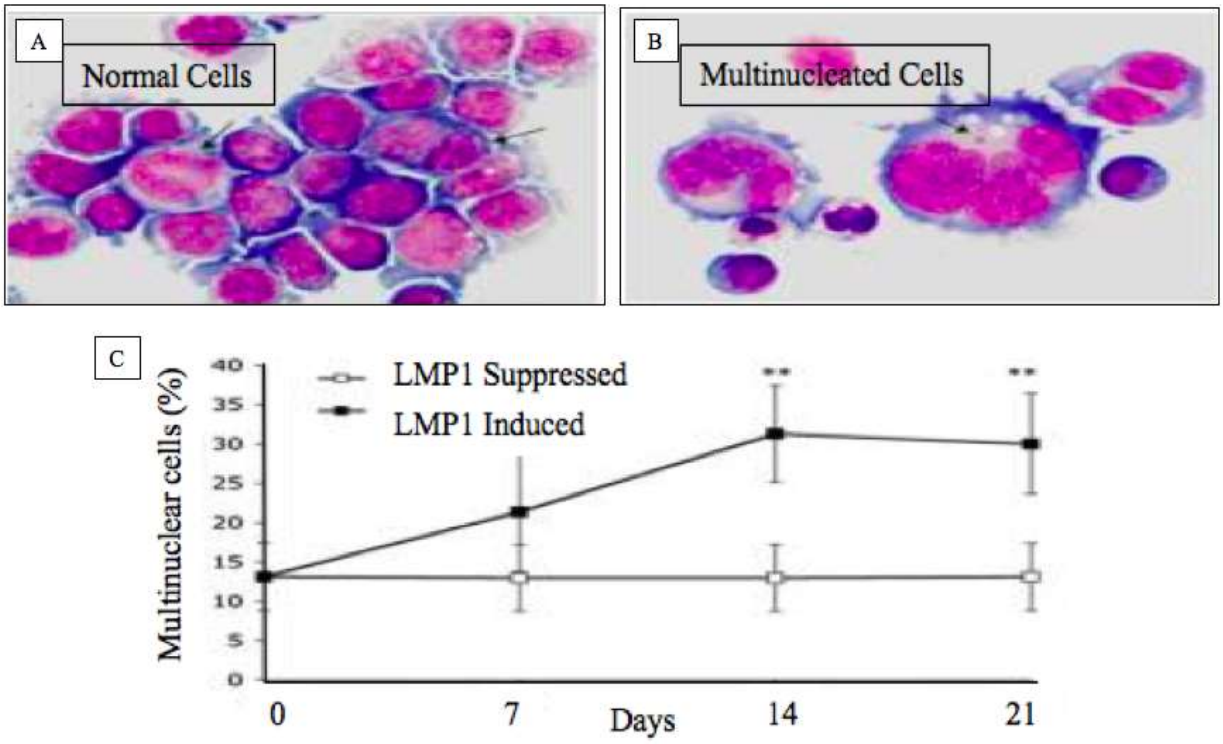


Figure 9: LMP1 expression in BJAB tTa LMP1 cell line.
A) cells without LMP1 are showing a stable percentage of bi or multinuclearity and B) cells expressing LMP1 showing a distinctive increase pattern of bi and multinuclearity, C) a graph representing the increase of multinuclearity over a period of 21 days (Adapted from Lajoie et al., 2015).

RS cells are characterized by telomeric disorganization, short telomeres and loss of total telomeres and same morphology can be observed in RS cells associated with LMP1. This led to the hypothesis that LMP1 indirectly interferes with the shelterin complex, thereby initiating the telomeric disruption and assisting in the formation of H and RS cells. To determine if LMP1 expression favored the formation of multinucleated cells, Lajoie et al., 2015 used a system a tet-off inducible LMP1 in a stable BJAB tTa LMP1 cell line lacking chromosomal translocations. LMP1-expressing and LMP1-suppressed BJAB tTa cells, as well as the parental BJAB tTa cell line not harboring the LMP1 oncogene, were analyzed over a period of 21 days for formation of multinucleated cells, (Figure 9 C) and the expression of key proteins of the shelterin complex at the transcriptional, translational, and topographic protein level. The results show that the telomeres and the shelterin complex are responsive to the expression of the LMP1 oncogene (Figure 10).

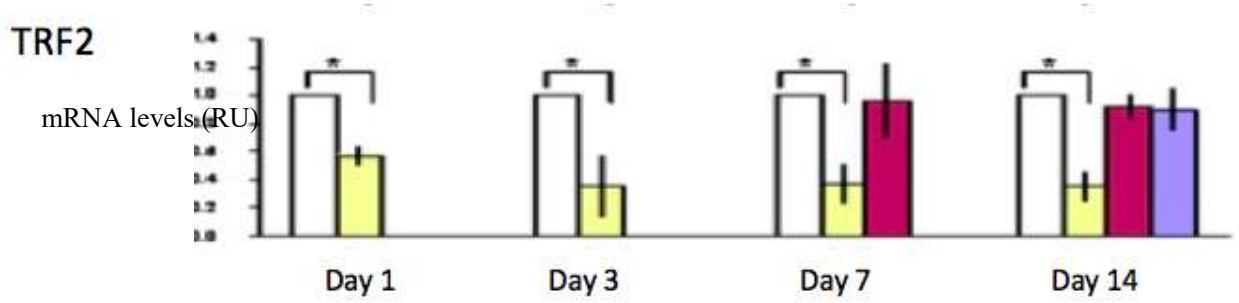


Figure 10: Quantitative reverse-transcriptase polymerase chain reaction analysis of the TRF2 mRNA upon LMP1 expression and suppression.

The TRF2 mRNA levels (in relative units) at days 1, 3, 7, and 14 were normalized with expression of 3 housekeeping genes to the LMP1-suppressed cells at each time point. Yellow bars represent TRF2 levels when LMP1 is expressed. Burgundy bars represent TRF2 levels after the expression on LMP1 was suppressed at day 3 and the blue bar when level of LMP1 suppressed at day 7 (Adapted from Lajoie et al., 2015).

The results also show that LMP1 exerts profound effects on shelterin protein expression, in particular the TRF1, TRF2, and POT1 proteins. TRF2 levels were the most affected when LMP1 was induced in the cells. Continuous LMP1 expression results in downregulation at the transcriptional level and is associated with a highly significant formation of multinucleated RS-like cells (Lajoie et al. 2015). Given that it was previously shown that deletion of TRF2 leads to chromosomal end fusions, bridges, and lagging chromosomes (van Steensel et al., 1998), it was suspected that it might be one of the key players for producing the multinucleated cells. To test this hypothesis, a myc-TRF2 driven by a CMV promoter was introduced into the BJABJ-tTA LMP1 cell line with and without the induction of LMP1 to verify if the multinucleation would still proceed. The independent expression of CMV-myc-TRF2 blocked the formation of multinucleated RS (Figure 11) thus showing that LMP1 appears to mimic the signaling pathway involved in the pathogenesis of the EBV-negative HL. Importantly, LMP1 induced changes in the expression of several shelterin proteins that act as gatekeepers for correct DNA replication telomeric integrity. These findings may eventually lead to the unraveling of the molecular pathogenesis of EBV-associated HL (Lajoie et al., 2015).

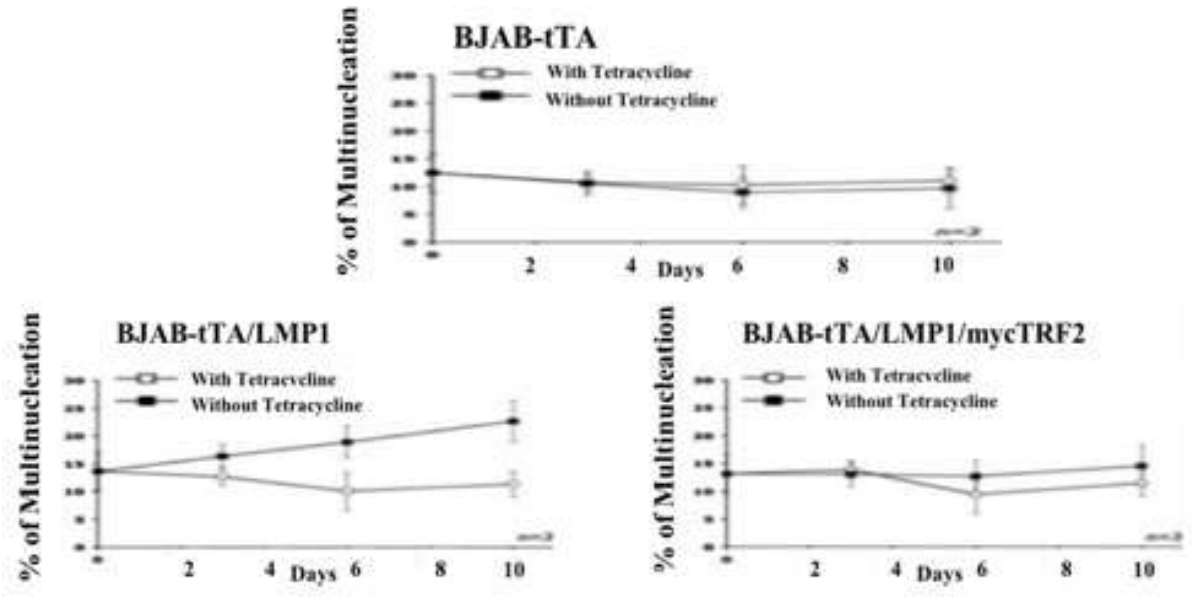


Figure 11: LMP1 increases multinucleation but expression of mycTRF2 eradicates multinuclearity in the presence of LMP1
Count of multinucleated cells in BJAB tTA and BJAB tTA LMP1 in the presence or absence of tetracycline over a period of 10 days (Adapted from Lajoie et al, 2015).

Hypothesis and Objectives

The ability of EBV viral proteins such as LMP1 to induce phenotypic changes when expressed individually in human B–cells lines implies that these viral proteins are key effectors in the immortalization process and changes of gene expression. No differences in clinical presentation are found between EBV negative and positive forms of HL, suggesting a common mechanism in pathogenesis. TRF2 appears to be a major target of LMP1 but is not the sole downstream effector of telomere dysfunction in this setting, therefore the hypothesis of this project is that LMP1 indirectly interferes with TRF2 at the promoter/gene level, since TRF2 levels are reversible upon removal of LMP1 (Figure 10, Figure 12). The objective of this project is to firstly determine if LMP1 indirectly interferes with the TRF2 promoter since LMP1 did not affect the regulation of pCMV-TRF2 when transfected into the BJAB tTa induced with LMP1. Secondly, we hope to identify a smaller region within the TRF2 promoter in which potential inhibitors or missing activators are affecting TRF2 levels.

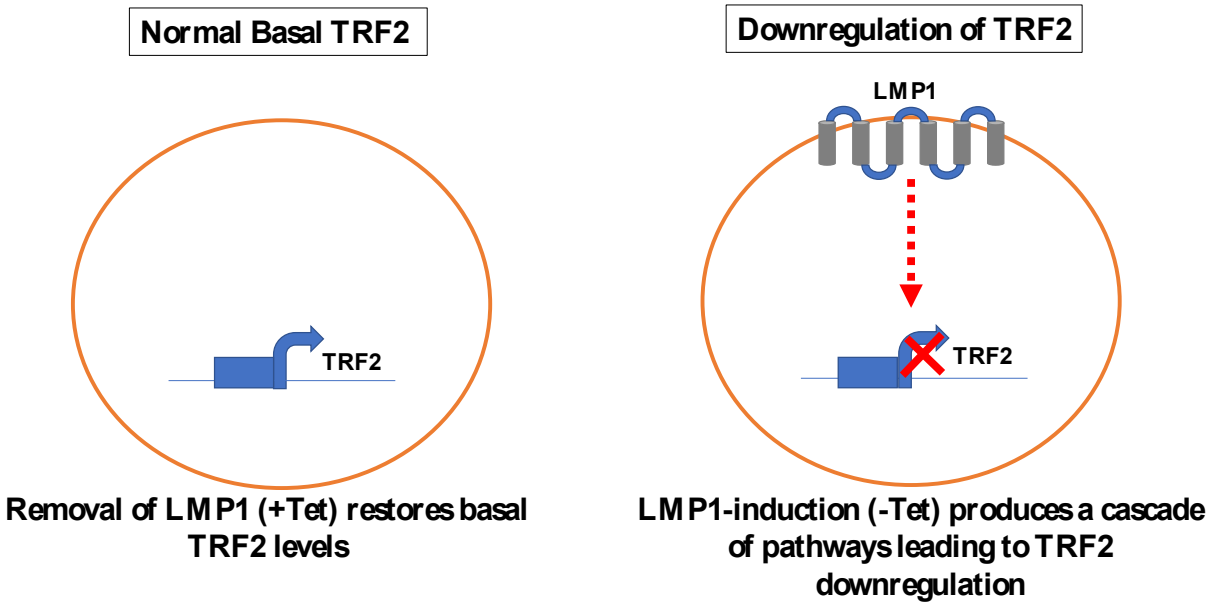


Figure 12: Schematic representation of LMP1 possibly interfering with TRF2 at its promoter region.
The hypothesis is that LMP1 indirectly downregulates TRF2 at the promoter, and when LMP1 is removed or stopped by the addition of Tetracycline, TRF2 basal levels are restored.

Materials and Methods

Cell Lines

The cell lines BJAB tTA and BJAB tTA LMP1 were obtained from Martin Rowe (Floettmann, Ward, Rickinson & Rowe, 1996). The BJAB cell line is an EBV-negative B lymphoma that shares many characteristics with typical Burkitt lymphoma except that it lacks the chromosomal translocation that is a marker for Burkitt Lymphoma diagnosis (Menezes, Leibold, Klein & Clements, 1975). Both cell lines have been integrated with a Tet-off system. The Tet-off system was made by adding a C-terminal portion of VP16 fused to the tetR-gene and is flanked upstream by a P_{CMV} and downstream by the SV40 poly(A) site. This tetR with the activation domain VP16 from HSV (Herpes simplex virus) generates a specific transactivator by binding to tetO sequences. Transactivation is prevented by the addition of the antibiotic tetracycline. When bound to TetO-placed upstream promoters, tTA (tetracycline-controlled transactivator) activates transcription (Gossen & Bujard, 1992).

BJAB tTA is the parental cell line of BJAB tTA LMP1, integrated with the plasmid pUHD15-1 that has been modified by insertion of a hygromycin-resistance gene to give rise to pJEF-3, which expresses the tTA system. The tTA system is derived from fusion of two genes: tetR and a VP16 transactivator that is continuously expressed from a cytomegalovirus immediate-early enhancer (CMV/IE) promoter on the pJEF3 plasmid. The TET repressor (tetR) protein is a transcription repressor from Escherichia coli (E. coli) and the vP16 domain transforms the tetR repressor into an activator (Floettmann et al., 1996). This will allow us control the gene expression by the addition or removal of Tetracycline. BJAB tTA LMP1 is the BJAB tTA cell line that has been integrated with not only pJEF-3 but also a plasmid called pJEF-6 containing a transactivator response element (TRE) that is located in front of the cytomegalovirus promoter (P_{CMV}) that controls the expression of the gene producing the LMP1 protein to produce stable clones requiring binding of tTA to a 5'regulatory region containing promoter (Kim et al., 2009). (Figure 13)

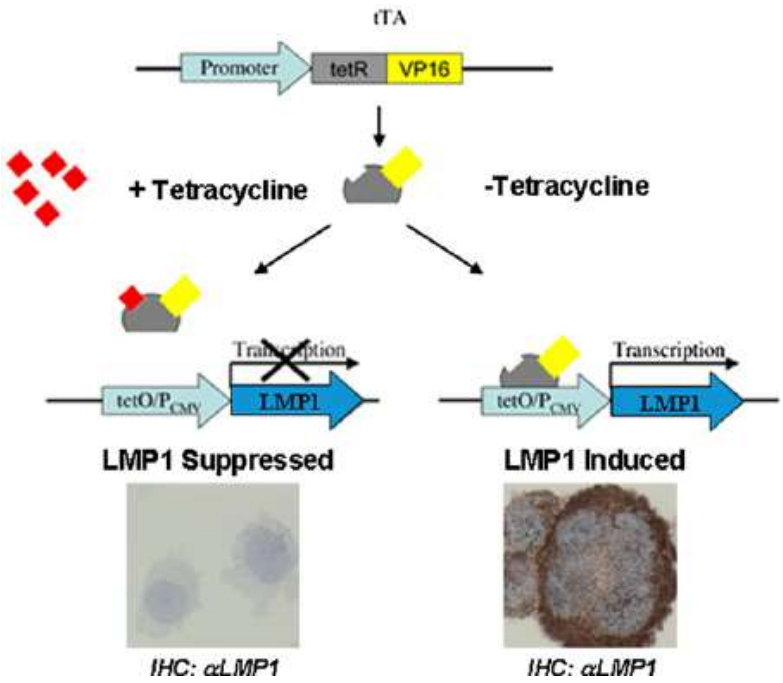


Figure 13: Tet-Off system.
The tTA binds specifically to the tet-Off (tetO) operator sequences upstream of the CMV promoter (P CMV). The tetO/P CMV represents the tetracycline response element (TRE) (Lajoie et al., 2015)

Manipulation of cell lines

Passage of Suspension Cultures

Replacement of growth medium is carried out every 3 days by directly diluting the cells in the culture flask for further growth or withdrawing a portion of the cells and diluting the remaining cells down to a seeding density appropriate for the cell line. This keeps the cell density in the range of continuous growth until used for a transfection experiment. Transfection requires 4 million cells per sample. Cells are always incubated at 37°C and 5% CO_2.

Cryopreservation of Mammalian cells

Mammalian cells are cryopreserved to avoid loss by contamination, to minimize genetic change, and to avoid aging or transformation of cell lines. Cells are counted in order to have 2 million cells per vial. The cells are centrifuged at 600 rpm for 5 minutes to pellet cells. Supernatant is then removed and the pellet is washed twice with 1X PBS. The cells are then resuspended in freezing medium composed of 90% RPMI-1640 and 10% DMSO and aliquoted into cryogenic

storage vials. The vials are placed at -20°C for 2-3 hours then placed at -80°C overnight before being transferred into a liquid nitrogen tank for storage (Freshney, 1994).

Thawing frozen Mammalian cells

Cell vials are removed from the liquid nitrogen, placed in a 37°C water bath for 15 minutes until thawed, and spun at 600 rpm for 5 minutes. The supernatant is removed and cells are resuspended in pre-heated medium (RPMI-1640 with FBS) and transferred into a 75ml vented cap flask (Freshney, 1994).

DNA Extraction Protocol

Cells were centrifuged 4000rpm/5min/4°C and the supernatant was removed. The pellet was rinsed twice with 1X PBS and was centrifuged at 4000rpm/5min/4°C. The pellet was resuspended in 100ul of Dit buffer (10mM Tris pH8.0+0.1% Polyoxyethylene 10 lauryl ether (decaethylene glycol)) to lyse cells. Proteinase K was added to a ratio of 1/100 (stock 10 mg/ml, final concentration of 100 μg/ml). The mix was incubated at 65°C for 2 hours to digest proteins. It was then incubated at 95°C for 15 minutes to inactivate the Proteinase K. The genomic DNA was stored at -20°C for future use.

Plasmid Construction

Using plasmids pGL3 and eGFP-N1, the luciferase and eGFP gene were isolated by digesting with XbaI and HindIII and purifying on 0.8% agarose gel. The vector fragment from pGL3 and eGFP insert fragment were ligated together and transformed into DH5α *E. coli*. This created plasmid JSTG-GFP (Figure 14) with a multi-cloning site for future promoter insertion and an ampicillin resistance gene (AMP). Primers to amplify the candidate promoter region of TRF2 were designed using NCBI reference NM_005652.4. The TRF2 promoter sequence was extracted and used to design specific primers using primer blast. The primers were checked against the whole human genome for specificity to the TRF2 promoter (see Table 7 for primers). The JSTG-GFP vector was digested with XhoI and HindIII to be able to ligate with the PCR product. All forward primers contained an XhoI restriction enzyme sequence and all reverse primers contained a HindIII sequence in order to facilitate ligation with the JSTG-GFP vector.

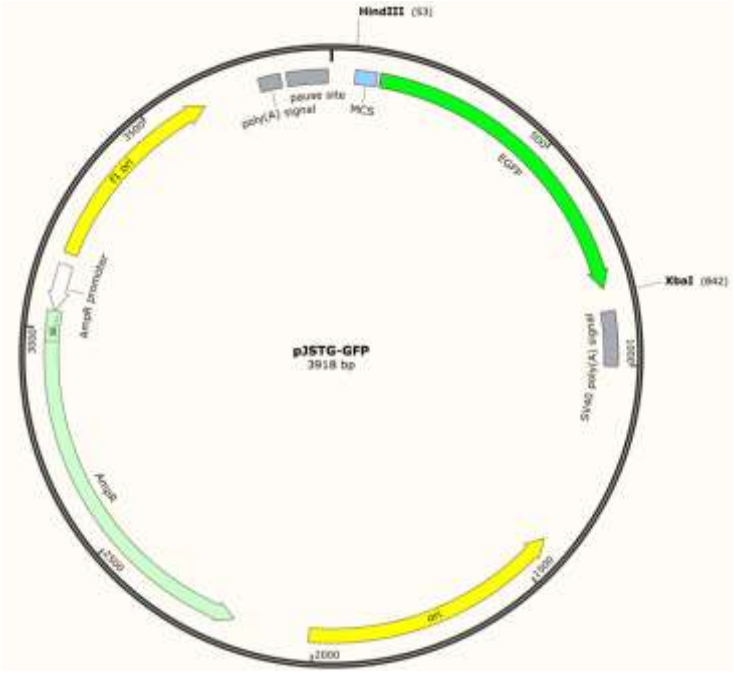


Figure 14: pJSTG-GFP plasmid construct.
A multi-cloning site and eGFP gene were integrated into a plasmid with the AMP gene using enzymes XbaI and HindIII.

PCR Mix:
2 µl of 10µM Forward primer
2 µl of 10µM Reverse primer
10 µl of 5X Q5 Buffer (NEB)
1 µl of 10 mM dNTPs
1.0 µl of Q5 polymerase (NEB)
1 µl DNA (~20 ng genomic DNA or 10 ng plasmid DNA)
33.0µl of ddH_2O
50 µl total volume

The following PCR conditions were used for all the reactions:
Step 1: 95°C for 2 minutes
Step 2: 95°C for 30 seconds
Step 3: 59°C for 30 seconds
Step 4: 72°C for 3 minutes
Repeat step from 2 to 4 for 30x
Step 5: 72°C for 10 minutes

Plasmid Construct Names	Length of TRF2 Promoter (bp)
TRF2-p2566	2566
TRF2-p1307	1307
TRF2-p1157	1157
TRF2-p974	974
TRF2-p821	821
TRF2-p737	737
TRF2-p680	680
TRF2-p614	614
TRF2-p484	484
TRF2-p400	400
TRF2-p300	300
TRF2-p250	250
TRF2-p200	200
TRF2-p150	150
TRF2-p140	140
TRF2-p125	125
TRF2-p115	115
TRF2-p100	100
TRF2-p50	50
TRF2-PLGFP	GFP-PL

Table 1: Name of plasmid construct with respective promoter length.

Electrophoresis Gel

PCR reactions were loaded on a 1X TAE 1% agarose gel and run at 100 V for 60 min. 1kb+ DNA ladder (Invitrogen) was loaded as a DNA marker. The gel was placed in an ethidium bromide bath for 30-60 seconds then run for 15 minutes and visualized using a UV transilluminator and photographed with digital software.

Ligation

A T4 DNA ligase kit (New England Biolabs, product M0202S) was used according to manufacturer's instructions for ligation to either clone an insert into a vector or circularize a vector after cloning. The samples were incubated for 30 min at RT. The reaction mixture was then stored at 4°C or used right away for transformation into *E. coli*.

Ligation of insert DNA into plasmid vector DNA

Reagents	Amount(µl)
Insert (3:1)	Varies
Vector	10-100 ng
T4 Ligation Buffer (10x)	4µl
T4 DNA Ligase (1U/µl)	1
Water	Add to make 20 µl
Total	20 µl

Table 2: Reagents and volume used for ligation of insert with vector.

Recircularization of linear DNA

Reagents	Amount(ul)
Vector	10-50 ng
T4 Ligation Buffer (10x)	10 µl
T4 DNA Ligase (1U/µl)	1 µl
Water	To 50 µl
Total	50 µl

Table 3: Reagents and volume used for circularization ligation.

Enzyme Digestion

All enzymes are obtained from NEB. All enzyme digestion reactions were incubated at 37°C for 1 hour.

Digestion Mix
1 µg DNA (purified plasmid DNA, PCR amplicons, vector plasmid)
1 µl of each Restriction Enzyme
5 µl of 10x Buffer
x µl dH_2O (to bring total volume to 30µL)

Klenow Treatment

This protocol blunts ends by 3'overhang removal or fill-in 3' resected (5' overhang) ends using DNA polymerase I, Large (Klenow) Fragment. Using 30ng purified digested plasmid, 2U Klenow, 1x buffer, 5mM dNTP, and water were mixed and incubated for 20 minutes at room temperature. The reaction was stopped by adding EDTA to a final concentration of 10mM and heating for 15 minutes at 75°C (NEB, 2015).

Bacteria Transformation

Bacterial transformation was performed using *E. coli* DH5α competent cells stored at -80°C and defrosted on ice. 50 µl was transformed with 5 µl of ligation product, pipetting up-and-down gently to mix, and incubated for 15 minutes on ice. The samples were heat shocked in a 42°C water bath for 90 seconds and then placed on ice for 2 minutes. 950 µl of LB broth was added and bacteria were incubated at 37°C for 1 hour with agitation. The bacteria were centrifuged for 2 min at 12000 rpm, 950ul of the supernatant was removed and bacteria was resuspended in the remaining 50 µl of LB and spread onto LB-agar plates containing a selection antibiotic such as ampicillin (100mg/ml). The plates incubated overnight at 37°C. Single colonies were selected and grown in 2 ml of LB + ampicillin overnight at 37°C. Plasmids were extracted using EZ-10 spin column plasmid DNA mini-prep kit (Biobasic) and aliquots were frozen at -20°C to be thawed at 4°C as needed for downstream applications.

BJAB tTA/LMP1 Transfection

Human suspension cell lines BJAB tTA/BJAB tTA LMP1 were transfected with a plasmid containing a portion of the TRF2-promoter linked to a GFP gene (Table 7) using either electroporation method or the nucleofection kit (Lonza 10064-154). The day before the transfection, the cells were washed twice with PBS 1x and incubated overnight with 0.8 µg/ µl tetracycline at 37°C and 5% CO_2. On the day of the transfection cells were counted using a hematocytometer and cell number was calculated as follows: Cell Concentration (cells/ml) = Dilution Factor (Total cell count/4) $x10^4$. 4.0 x 10^6 cells per sample were used to perform the transfection. Cells were washed twice 1x PBS, suspended in RPM1+FBS without tetracycline and incubated for 4 hours to remove all traces of tetracycline in order to induce LMP1 expression.

When the electroporation procedure was used, it was performed as follows: after 4 hours of incubation, the cells were washed twice with PBS to remove RPMI+FBS and resuspended in 100 µl per sample of RPM1 without FBS to have 1 million cells per 100 µl. One hundred microliters of the solution with cells was transferred into 1.5 ml Eppendorf tubes with their respective amount of plasmid (5µg) required for transfection. Once samples were homogenized with the

plasmid, the mixture was added to the 100 µl cuvettes and transfected using a cuvette of 0.2 cm with the program conditions of 975 uF, 200 ohms, 150 V to pulse. After transfection, 100 µl of RPMI+FBS was added into each cuvette containing cells, then the 200 µl was transferred into 1.5 ml Eppendorfs where it was split: 100 µl was pipetted into 6 well plate with 5ml RPM1+FBS and the other 100 µl was pipetted into 5 ml RPM1+FBS+Tetracycline. The cells were incubated for 72 hrs. After 24 hrs, the media was changed by spinning the cells for 5 min at 600 rpm, removing the supernatant, and adding fresh media. Cell were harvested after 72 hrs, spun for 5 min at 1000 rpm in 15 ml plastic tubes, washed once with 1x PBS and transferred into 1.5 ml Eppendorf tubes. The cells were washed another time with 1x PBS and the pellet was left in 50 µl of PBS (for easier resuspension of cells for Protein Lowry or RNA extraction) and were preserved at -20°C.

If the nucleofection technique was used with Lonza nucleofection kit: after 4 hours, the cells were washed twice with PBS to remove RPMI+FBS and were resuspended in 100 ul per sample of reagent to have 4 million cells per 100 µl. One hundred microliters of the solution with cells was transferred into 1.5ml Eppendorfs with their respective amount of plasmid (3µg) required for transfection. Once samples were homogenized with plasmid, the mixture was added to the 100 µl cuvettes and transfected using the program provided by the company. After transfection, 100 µl RPMI+FBS was added to each respective cuvette containing cells, then the 200 µl was transferred into 1.5 ml Eppendorfs, where it was split: 100ul was pipetted into a 6 well plate with 5 ml RPM1+FBS and the other 100ul was pipetted into 5 ml RPM1+FBS+tetracycline. The cells were incubated for 72 hrs. After 24 hrs, the media was changed by spinning the cells for 5 min at 600 rpm, removing the supernatant, and adding fresh media. Cells were harvested after 72 hrs, spun for 5 min at 1000 rpm in 15 ml plastic tubes, washed once with 1x PBS and transferred into 1.5 ml Eppendorfs. They were then washed another time with 1x PBS and the pellet was left in 50 ul of PBS (for easier resuspension of cells for Protein Lowry or RNA extraction) and was preserved at -20°C.

Flow Cytometry

After the 72 hours transfection, the cells were aliquoted into a 15 ml tube, spun for 5 min at 600 rpm then washed twice with 1x PBS and resuspended in 1 ml of fresh cold PBS and transferred

into the flow cytometry tubes. The FACS DIVA system was used according to manufacturer's instructions with the following parameters: FSC 165 SSC 245 GFP 385 intensity).

Fluorescent Microscopy

To determine if the transfection was successful, after 72 hrs, the cells were left in the 6 well plates and centrifuged at 1000 rpm. Once the cells were spun down, the 6 well plates were brought to the microscope. A pre-established program by Sheriff Lab was used for analysis.

Western Blot

The pellet previously harvested or stored at -20°C was resuspended in 50 µl of 2X Laemmli loading buffer with DDT added (62.5 mM Tris pH 6.8, 25% v/v glycerol, 2% m/v SDS, 0.01% m/v bromophenol blue, 10% v/v 1M DDT and 50 µl of water. The samples were sonicated for 10 min at 100 V in an ice bath (1 min sonication with 25 sec pause). The samples were boiled for 5 minutes at 100°C in order to denature the proteins. The amount of protein per microliter was quantified using Protein Lowry colorimetry, which is a widely used method to estimate the amount of proteins within a biological sample, (Lowry, Rosenbrough, Farr & Randall, 1951) and 20 µg of each sample was then analyzed on a 10% SDS-PAGE gel. For all Western blots, samples were migrated for 1 hour 40 min at 100 V. The proteins where then transferred to Hybond-C nitrocellulose membrane (GE Healthcare) using a BioRad mini Trans-blot Cell apparatus for 75 minutes at 100 V. Membranes were stained with Ponceau S in order to verify the quality of the proteins and transfer efficacy. The membranes were blocked with 5% m/v powdered milk dissolved in 1X PBS-T for 30 minutes. They were then washed 3 times for 5 minutes each in PBS-T with agitation at room temperature. Membranes were incubated with rotation overnight at 4°C with antibody (Table 4). They were then washed 3 times for 5 minutes in 1X PBS-T at room temperature with agitation. Membranes were incubated for 1 hour at room temperature with the secondary antibody (Table 5). After this incubation, the membranes were washed again 3 times for 5 minutes with 1X PBS-T. Enhanced 32 chemiluminescence (ECL) was used to visualize the results with an ImageQuant LAS 4000 (GE Healthcare).

Primary Antibodies used for Western Blot

Antibody	Company	Dilution	Expected Size	Secondary Antibody
Anti-eGFP	Roche	1:1000	26kDa	HRP-conjugated anti-mouse
Anti-Actin	Sigma Aldrich	1:1000	45kDa	HRP-conjugated anti-rabbit
Anti-TRF2	IMGENEX	1:1000	60kDa	HRP-conjugated anti-mouse
Anti-Tubulin	Cell Signaling	1:1000	65kDa	HRP-conjugated anti-rabbit
Anti-LMP1	Sigma Aldrich	1:200	70kDa	HRP-conjugated anti-mouse

Table 4: Primary antibodies with their properties for Western Blot usage.

Secondary Antibodies used for Western Blot

Secondary Antibody	Company	Dilution
Mouse	Cell Signaling	1:5000
Rabbit	GE Healthcare Life Sciences	1:5000

Table 5: Secondary antibodies and dilution.

RNA Analysis

Total cellular RNA was extracted by Trizol reagent and RNA extraction kit (Zymo Research, Direct-zol kit R2050, R2051) and purity of each sample was measured by optical density using a Nanodrop spectrophotometer. Total RNA (1 µg) was used to generate cDNA using the Qiagen QuantiTect Reverse Transcriptase kit following the manufacturer's instructions. Real-time PCR was performed with a hot start step 95°C for 10 min followed by 40 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 30 s using iQ^{TM} SYBR Green Supermix (Bio-Rad Laboratories Inc.) on Real Time PCR System (Applied Biosystems).

Nuclear Extraction Protocol

Cells were incubated at 37°C and 5% CO_2 until a density of 1x10^{9}cells/culture was obtained. The cells were spun for 10 min at 2000 rpm in 50 ml plastic Falcon tubes at room temperature. Supernatant was removed and the cells were washed twice with cold 1x PBS for 5min at 2000 rpm and 4°C. The cells were then resuspended in 5 volumes (compared to pellet volume) of autoclaved Buffer A (10 mM Hepes pH 7.9, 1.5mM $MgCl_2$, 10 mM KCl and 0.5 mM DTT). The pellets with buffer were incubated for 10min on ice, centrifuged for 5 min at 2000 rpm at 4°C

and supernatant was removed. The pellet was then resuspended in 2 volumes of Buffer A and transferred into a 7 ml glass pestle Dounce homogenizer Type A with 15 strokes to lyse the cells. The mixture was then transferred to 15 ml Falcon tube and centrifuged for 10 min at 2000 RPM at 4°C to precipitate the nuclei. Supernatant containing the cytoplasmic mixture was removed and stored at -80°C for other experiments and the pellet was resuspended in 2 volumes (compared to nucleus pellet volume) of the Buffer A and centrifuged for 20 min at 17000 RPM (25000g) at 4°C (rotor SS34, Sorvall). The supernatant was removed and the pellet resuspended in 2 volumes of Buffer C (20mM Hepes pH 7.9, 25% Glycerol, 420 mM KCl, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSF). The mixture was homogenized using the 7ml glass pestle dounce homogenizer Type B with 10 strokes on ice. The mixture was incubated on ice for 30 min with occasional agitation. The pellet mixture was then centrifuged for 30 min at 17000 RPM. The supernatant was transferred into a dialysis membrane (Spectra/Por 7 Dialysis Membrane MWCO 10Kd, #132119) in 1L Buffer D (20 mM Hepes pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT and 0.5 mM PMSF) for 1h-2h at 4°C lightly mixing with a stir bar. The dialysis buffer was changed an additional time and left overnight at 4°C. The next day, buffer D was changed and fresh buffer was added for 1h at 4°C. The samples were then centrifuged for 20 min at 17000 RPM and aliquoted on ice. Sample nuclear proteins were then measured using the Bradford technique (Bio-Rad Solutions), using a BSA standard curve and measured at 595 nm using a spectrometer.

5’ γ-ATP32 labelling

For the electrophoretic mobility shift assay (EMSA), the single stranded oligonucleotides were 5’ labelled with γ-ATP^{32}. A total reaction volume of 20µl was comprised of 5µl of the 1uM oligonucleotides, 2µl 10x PNK Buffer, 8µl sterile water, 4µl Gamma-ATP^{32} and 1µl of T4 PNK (10U/ml). The mixture was incubated at 37°C for 45 min and the reaction was stopped by heating at 65°C for 10 min. For the double stranded portion of the experiment, radiolabelled oligonucleotide were annealed together as described below.

Annealing of Oligonucleotides

In an Eppendorf, 22.5 µl of sense and 22.5 µl of antisense single-stranded oligonucleotides were mixed together with 5 µl of T4 DNA ligase buffer. The mixture was heated for 2 min at 95°C,

and immediately the Eppendorf (held by a float) was put into a beaker of 1L of boiling water. The mixture was gradually cooled by leaving it at room temperature. Once the mixture reached room temperature, the double-stranded oligonucleotide was stored at −20°C.

EMSA

A 4.5% native polyacrylamide gel was made fresh the day of the experiment. The gel was prepared using a medium plate with an 18 well spacer comb. The gel was comprised of 6 ml of acrylamide (30% 29:1), 4 ml 5x TBE Buffer (0.45 M Tris–borate, 10 mM EDTA), 2 ml glycerol (20%, vol/vol), 400µl 10% APS, 40ul 100% TEMED and 28ml sterile water (total volume 40 ml). Once poured, the gel was left to solidify for 30 min-1 hour. The gel was pre-run at 150V in 0.5x TBE with 2% glycerol buffer for 1 hour at 4°C. For EMSA sample preparation, all solutions were kept on ice. Each sample was brought to a final volume of 20µl, containing 4µl 5x binding buffer (100 mM HEPES, 25 mM $MgCl_2$, 40% glycerol, 500 mM KCl in distilled water) 0.4 µl 10 mg/ml BSA, 0.2 µl 10 mg/ml Poly(dI-dC)·Poly(dI-dC), 0.2 µl dsDNA OligoB (90 µM stock), and 10 µg of nuclear extract (volume depending on the concentration of the nuclear extract). The mixture was incubated on ice for 5 min. Afterwards, 1 µl of labelled dsDNA oligonucleotide was added, mixed and incubated at room temperature for 30 min. Samples were loaded using 2 µl DNA dye. Once loaded, the gel was run for 4hours at 150 V in a 4°C fridge. Once running was completed, the gel was transferred to Whatman paper and vacuum pump dried at 80°C. The gel was exposed on radiography film (GE Healthcare, product 28906845) for 4h and stored at -80°C.

DNA Affinity Chromatography (Pull-Down) Assay for DNA-Protein Affinity

DNA Affinity Chromatography was performed using a 30 bp sequence from a region of 100-130bp of the TRF2 promoter. A 5’ biotin labelled sequence was added to the 30 bp sequence on the reverse strand of the DNA with an additional Teg tag (ordered from Integrated DNA Technologies). The assay was performed as follows: Step 1: Binding of oligos biotinylated on Dynabeads M-280 Streptavidin (catalog number: 11206D). For each sample, 100 µl of magnetic beads was aliquoted into a silicone Eppendorf tube and put on a magnetic particle concentrator (Invitrogen, product # 123.21D) for 2 min. The supernatant was removed and beads were washed 3 times with 100 µl of B&W 1x buffer (5mM Tris-HCl pH7.5, 0.5 mM EDTA and

1M NaCl). The beads were resuspended in 200 µl of B&W 2x buffer (10 mM Tris-HCl pH 7.5, 1mM EDTA, and 2M NaCl). To each tube, 188 µl nano water with 12 µl of biotinylated oligo (50 µM:600pmol/µl/load) was added for a total volume of 400 µl. The tubes were agitated by flicking every couple of minutes to resuspend the beads for a total time of 15 min. The tube was put onto the magnet bar for 2 min to collect the beads and the supernatant was removed. The beads were washed three times with B&W 1x buffer and wash once with 200 µl of BB 1x buffer (20mM HEPES ph7.9, 100mM KCl, 5mM $MgCl_2$, 8% Glycerol). After the washes, the buffer was removed and 100ul of BB 1x buffer was added. The tubes were then placed on ice. Step 2: Addition of nuclear extract. A total of 900 ul of Master Mix NE reaction was added to the beads and oligo that were resuspended in 100ul of BB1x buffer. For one sample, the final volume of the experiment was 1 ml. With the 100 µl of BB 1x already added to the magnetic beads with biotin oligonucleotide, a master mix was prepared with 300 µl of BB5x buffer (100mM HEPES ph7.9, 500mM KCl, 25mM $MgCl_2$, 40% Glycerol), 30 µl of acetylated BSA (10 mg/µl), 15 µl of ds Oligo B (90 µM), 15 µl of Poly dIdC 10 mg/µl), and 1.5 mg of nuclear extract (varies depending on the concentration). All products were added together and incubated for 3 min. 0.05% CHAPS was added and left for 2 min of interaction with the mix. After 5 min, the nuclear extract mix was added to the beads and oligonucleotide. The nuclear extract and dsDNA were left to interact for 15 min. Samples were placed on the magnet bar for 2 min to collect the beads and then washed according to the time conditions in table 6.

Samples	**#1**	**#2**	**#3**
	Time in Minutes		
Beads	0	1	2
Wash	5	6	7
500 µl BB 1x + 0.05% CHAPS	7	8	9
Beads	12	13	14
Remove then wash with 500 µl BB 1x (0.01% NP40)	14	15	16
Beads	19	20	21
Remove then wash with 500µl BB 1x (0.01% NP40)	21	22	23
Beads	26	27	28
Remove then wash with 500 µl BB 1x (0.01% NP40)	28	29	30
Beads	33	34	35

Wash BB 1x Keep 100 µl for Silver Stain	35	36	37
Beads	40	41	42

Table 6: Example of a time course for the pull-down for 3 samples. Every sample was separated by 1 min for easier manipulation.

Mass Spectrometry

Mass spectrometry (MS) analysis of proteins measures the mass-to-charge ratio of ions to identify and quantify molecules in simple and complex mixtures. Samples were prepared using protocol "On-Beads Trypsin Digestion Boisvert Lab, May 2014". This method allows better recovery of the immunoprecipitated proteins, at the cost of increased contamination with antibody and increased complexity of the sample. The steps are as follows: step 1 - digestion including reduction and alkylation. This step is required to prevent disulfide bonds between peptides, thus increasing peptide coverage. However, there are several reports indicating that antibodies which have not been reduced may be partially resistant to trypsin, which could help reduce the amount of antibodies in the sample. 1) Wash beads 5 times with 20 mM ammonium bicarbonate. 2) After the final wash step, add enough 10 mM DTT in 20 mM ammonium bicarbonate to the beads for complete immersion, mix and incubate at 60°C for 30 minutes. 3) After cooling, add an equal volume of 15 mM iodoacetamide in 20 mM ammonium bicarbonate to the DTT/bead suspension, mix and incubate in the dark for 1 hour. 4) Add 1M DTT to increase the concentration to 15 mM to quench the iodoacetamide and wait 10 minutes. 5) Add 50 ng of trypsin to the beads. 6) Incubate at 37°C for 5 hours to overnight. 7) Stop trypsin digestion by acidifying to a final concentration of 1% formic acid. 8) Harvest the supernatant and transfer to a clean, protein lo-bind tube. 9) Resuspend beads in 60% ACN – 0.1 FA at room temperature for 5 min. 10) Transfer this second supernatant with the first supernatant. 11) Dry samples in the speed vac 3 hours at 65°C. 12) Resuspend in 20 µl sample buffer (0.1% TFA) to desalt on ZipTip. Samples were then measured by nanodrop. The samples were analyzed by Dominique Levesque from Francois-Michel Boisvert Lab.

Primers used in this study

Primers Name	Sequence (5’ -3’)
Plasmid Construct Primers	
R13 (Starts at the ATG of TRF2 promoter)	GACTAAGCTTGATAGAAACAGCGTTCCGAGCC
F2GC	GACACTCGAGATGGCATGAACCCAGAGGC
F12	GACACTCGAGACTCATCCAGGACCTGCCTA
F24	GACACTCGAGGTGATTACAAAGAGTATATGTCTAGCGGCTGG
F25	GACACTCGAGCCTCTGTGTATGCTCTCTTCCACAATTTTCA
F27	GACACTCGAGGTATCACACTGGTCTTTGCTTTATAATTACAG
F28	GACACTCGAGGCGATCTCGGCTCACCGCAAC
ALL-RW400	GACACTCGAGACAGGCATGCGCCACCAC
ALL-RW300	GACACTCGAGTCCACCCGCCTCAGCC
ALL-RW250	GACACTCGAGCGATCCCGGCCTGTTTTTCAGT
ALL-RW200	GACACTCGAGTAGTAGCTGTTTTCTGTAAATTCGG
ALL-RW150	GACACTCGAGGGGTTCACCGGCCTGGG
ALL-RW140	GACACTCGAGCCTGGGCGGTGTGGACT
ALL-RW125	GACACTCGAACTTCAAGCCCCAGGAGGC
ALL-RW115	GACACTCGACCAGGAGGCATTGCGGC
ALL-RW100	GACACTCGAGGCCGGCACATCGGGAACTAC
ALL-RW50	GACACTCGAGTCCTTTCCCTCCCAGAAGCC
ALL-RWR3	GACACTCGAG CAGGTGGCACTTTTCGGGGAAATG
RT-qPCR Primers	
ACTB For	ATGTCTCGCTCCGTGGCCTTA
ACTB Rev	ATCTTGGGCTGTGACAAAGTC
GAPDH For	GCGGGAAATCGTGCGTGACATT
GAPDH Rev	GATGGAGTTGAAGGTAGTTTCGTG
B2M For	CAACCAATAGAGTCCACCAGT
B2M Rev	TCTTCAGAAGAGACCTTCTCT
LMP1 For	CAACCAATAGAGTCCACCAGT
LMP1 Rev	TCTTCAGAAGAGACCTTCTCT
TRF2 For	GTACCCAAAGGCAAGTGGAA
TRF2 Rev	TGACCCACTCGCTTTCTTCT
GFP1 For	AGATCCGCCACAACATCGAG
GFP1 Rev	GTCCATGCCGAGAGTGATCC
GFP2 For	CTACCCCGACCACATGAAGC
GFP2 Rev	AAGTCGATGCCCTTCAGCTC
GFP16 For	CAACCACTACCTGAGCACCC
GFP16 Rev	GTCCATGCCGAGAGTGATCC
GFP18 For	CGACCACTACCAGCAGAACAC
GFP18 Rev	GTCTTTGCTCAGGGCGGACT
DNA Affinity Chromatography Primers (Pull-Down)	
100-130 F	TGTGGACTTCAAGCCCCAGGAGGCATTGCG
100-130 R Anti-Sense	Biotin-TEG-CGCAATGCCTCCTGGGGCTTGAAGTCCACA

Oligo B+ (competitor)	GATCCGGAGTACTTCAAGAACG
Oligo B- (competitor)	GATCCGTTCTTGAAGTACTCCG
EMSA Primers	
100-130 F	TGTGGACTTCAAGCCCCAGGAGGCATTGCG
100-130 R	CGCAATGCCTCCTGGGGCTTGAAGTCCACA

Table 7: Primers used in this study

Results-Chapter 1

TRF2 promoter plasmid constructions and analyses

Preamble

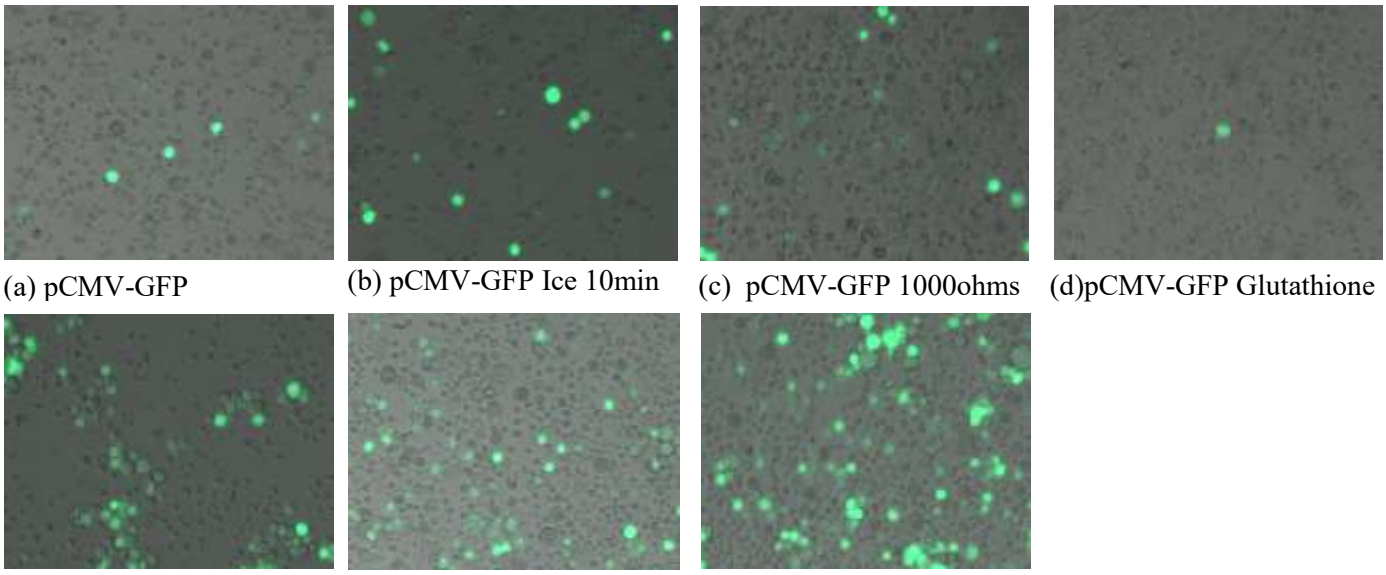
LMP1 has been shown to affect gene activation/inhibition pathways, leading to shorter telomeres with a decrease of TRF2 levels. The mechanism by which it does this has been linked to those involved in the pathogenesis of EBV-negative Hodgkin Lymphoma (Knecht et al., 2009). Understanding the role of LMP1 in the production of Hodgkin's and Reed-Sternberg cells will provide insight into the mechanisms through which B lymphocyte cells become Reed-Sternberg cells with/without the presence of EBV. To investigate our hypothesis, a BJAB tTA LMP1 cell line was used to control the expression of LMP1 by the addition or removal of tetracycline. To determine if the LMP1 pathway interacts with TRF2 at its promoter region, plasmids containing different TRF2 promoter lengths were transfected into the BJAB tTA LMP1 cell line and analyzed by Western blot, flow cytometry, RT-qPCR and/or microscopy to see whether differences in GFP protein or GFP RNA levels could be observed.

pTRF2-GFP Construct and Transfection Optimization

The first objective of this project was to construct a plasmid containing an ampicillin resistance (AMP) gene and the multi-cloning site (MCS) in front of a GFP reporter gene. A brief description of the process will be given here, with full details explained in the materials and methods section. Plasmid pGL3 was used for the backbone as it already contained the AMP gene and MCS but had a luciferase gene that was ultimately replaced by a GFP gene from plasmid pEGFP-N1. Using restriction enzymes XbaI and HindIII, the luciferase gene was removed in order to use the plasmid as a vector and the GFP gene was inserted in its place. The vector and insert were placed on a 0.8% TAE gel, purified using a gel purification kit (Biobasic), ligated and transformed into *E. coli*. The plasmid was then extracted from *E. coli* and sent for sequencing (https://www.sequences.ulaval.ca). This new plasmid construct was named PL-GFP. The next step was to clone different sizes of the TRF2 promoter that were obtained by extracting the genomic DNA from the cell line BJAB tTA LMP1 and using it for PCR amplification of the

desired primer region. PCR was performed with forward primers F12, F2GC and reverse primers R13. The forward primers have an XhoI site and the reverse primers have a HindIII site. The plasmid called PL-GFP and the PCR products were digested with XhoI and HindIII enzymes. They were then ligated and transformed into *E. coli*. The plasmid was then extracted from *E. coli* and sent for sequencing (https://www.sequences.ulaval.ca). The plasmid using F12-R13 was named p2566-GFP and the plasmid that used F2GC-R13 was called p1307-GFP. These plasmid constructs were then analyzed using microscopy, flow cytometry, and Western blot.

To determine the efficiency of the transfection, fluorescence microscopy was used for analysis. Electroporation was used with different conditions as well as nucleofection to determine which technique produced the better percentage of cells transfected with the GFP-containing plasmid. For optimization purposes, the pCMV-GFP plasmid was used since this plasmid contains a strong viral promoter and therefore should express maximum GFP proteins. The pCMV-GFP plasmid was transfected using different conditions, as seen in Figure 15 and 16: (a) condition is the normal condition mentioned previously in the materials and methods, (b) after the transfection, the cells were placed on ice for 10min (Potter & Heller, 2003), (c) the ohms were changed for 1000ohms, (d) 5mM of glutathione was added before the transfection to help prevent leakage of the cytoplasmic components and protection of membrane against oxidation (van den Hoff, Moorman & Lamers, 1992) (e) nucleofection using program called DN-100, (f) nucleofection using program called DS-137 and (g) nucleofection using program called EN-150 (Lonza Nucleofector Technology). By taking random picture frames, the number of transfected cells was counted to determine which experiment yielded a higher percentage of transfected cells.



(a) pCMV-GFP (b) pCMV-GFP Ice 10min (c) pCMV-GFP 1000ohms (d)pCMV-GFP Glutathione

(e) pCMV-GFP DN-100 (f) pCMV-GFP DS-137 (g) pCMV-GFP EN-150

Figure 15: Fluorescence microscopy.
pCMV-GFP (positive control) was transfected using different programs and cells were imaged using fluorescence microscopy to find the optimal transfection conditions for further experimentation.

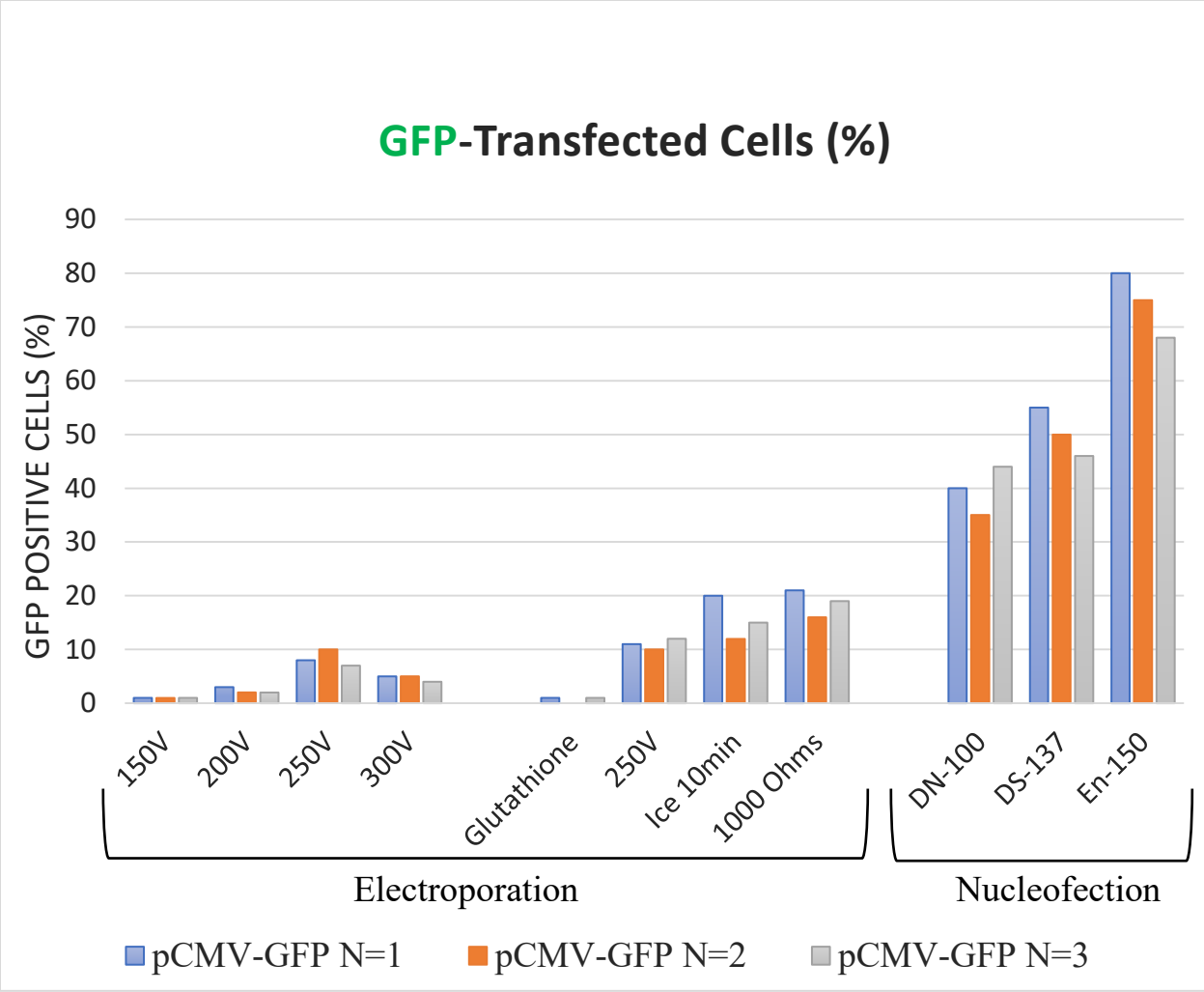


Figure 16: Graph depicting the transfection analysis of pCMV-GFP using BJAB tTA cell line.
Electroporation vs Nucleofection after 72h of transfection**.** Electroporation yielded 10-20% of positive cells with GFP while Nucleofection yielded a positive transfection of 40-80%. Nucleofection using program En-150 provided 80% GFP positive cells.

Western Blot Analysis of promoter regulation via GFP protein levels

Once the program and technique producing the greatest transfection GFP yield were determined, the next step was to verify the protein levels by Western blot. As mentioned previously, LMP1 indirectly interferes with the RNA levels of TRF2 (Lajoie et al., 2015). By integrating a portion of the TRF2 promoter linked to a GFP reporter gene, we wanted to determine if LMP1 causes a downregulation due to its influence on the promoter region of TRF2. A transfection was performed using nucleofection with the program EN-150, that was previously

determined to be best for future experiments. The experiment consisted of integrating the plasmid within the BJAB tTA LMP1 cell line by using the negative control called PL-GFP (promoterless plasmid), p1307-GFP, p2566-GFP and positive control pCMV-GFP. Figure 17 represents the results of the experiment and the protein analysis via Western blot. The proteins were quantified using the protein Lowry protocol. The membrane was probed with anti-LMP1, anti-GFP, and anti-actin as a housekeeping gene loading control.

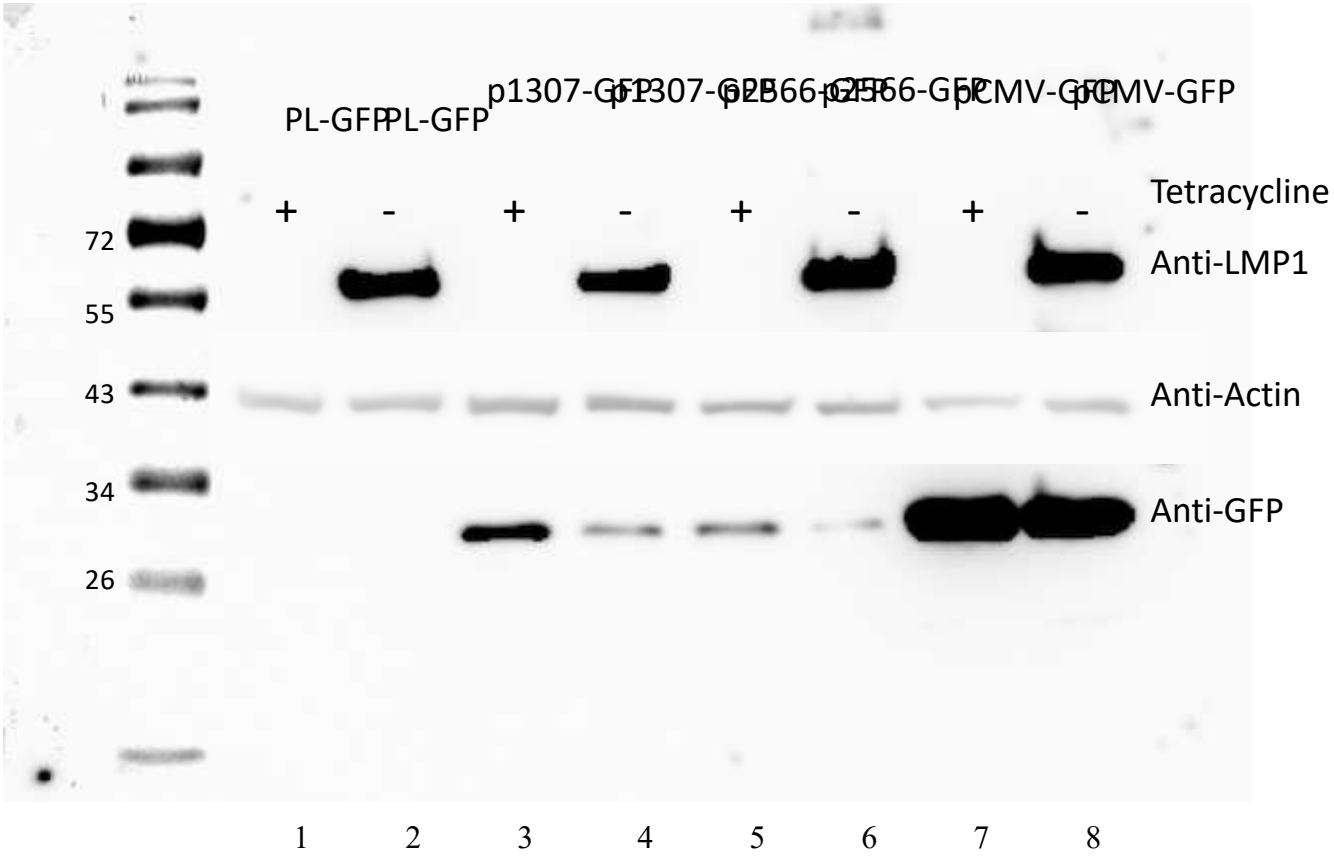


Figure 17: Western blot of TRF2 promoter analysis.
Lanes 1-8 represent the different promoters used in front of the GFP reporter gene as indicated on top of lanes.

The loading control anti-actin demonstrates that the proteins were evenly loaded. Lane 1 and lane 2 are the negative control, in PL-GFP Lanes 3-4 are the p1307-GFP construct. LMP1 is absent in lane 3 and present in lane 4. A downregulation of the GFP levels can be observed when comparing lane 3 to lane 4. Lanes 5-6 are the p2566-GFP construct, with LMP1 absent in lane 5 and present in lane 6. A downregulation of GFP is again observed when comparing lane 5 to lane

6. Lanes 7 (LMP1 absent) and 8 (LMP1 present) are the positive control pCMV-GFP. There is no apparent change of GFP protein levels in the positive control.

Analysis of GFP through Flow Cytometry

This technique is based on biophysical technology in fluorescent cell counting. It measures the cells one by one, in a rapid flowing fluid where the cells are passed through an optical system that excites the stained cells with a light source thus producing a fluorescence emission. This information is processed through the electronics component of the flow cytometer (ThermoFisher Scientific). Using samples from the previous transfection as seen in the Western blot above, 1 ml was taken for flow cytometry analysis. The goal of this experiment was to see a distinguishable horizontal shift between samples with LMP1 versus samples with no LMP1. The objective was to determine if LMP1 indirectly downregulates TRF2 via its promoter. It was expected that in samples where LMP1 is present, the GFP would be shifted to the left and in samples without LMP1, GFP expression would shift to the right. Figure 18 represents the samples transfected with negative control PL-GFP. Figure 19 contains both PL-GFP (negative control) and CMV-GFP (positive control). CMV is a strong viral promoter and therefore a shift to the far right would be expected due to high expression of GFP protein. In the constructs used in this study, TRF2 does not have a strong promoter and therefore it was expected that the results would be found between the PL-GFP and the CMV-GFP. Figure 20 is the experiment with p2566-GFP, the purple line is without LMP1 and the blue line is with LMP1. A very minimal shift with and without LMP1 was observed. Figure 21 is the experiment with p1307-GFP, the purple line is without LMP1 and the blue line is with LMP1. Again, a very minimal shift with and without LMP1 was observed. These shifts within the data presented in Figure 19-22 are not significant enough to conclusively determine that LMP1 downregulates TRF2 at the promoter level. Therefore, no further experiments were analyzed using flow cytometry as it is not sensitive enough to detect minimal changes.

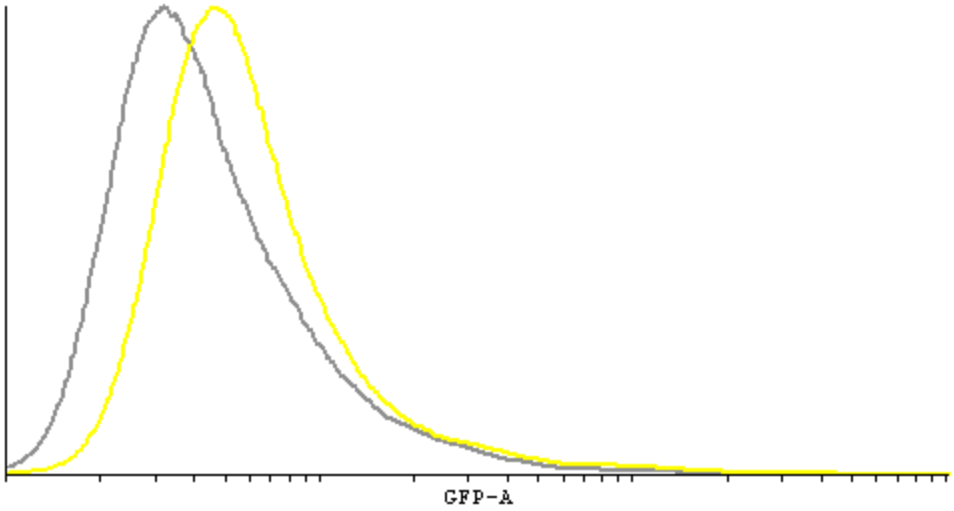


Figure 18: FACS analysis of PL-GFP.
This graph represents the negative control in the FACS experiments. The yellow line represents cells with no LMP1 induction (+Tet) and the grey line represents cells with LMP1 induction (-Tet).

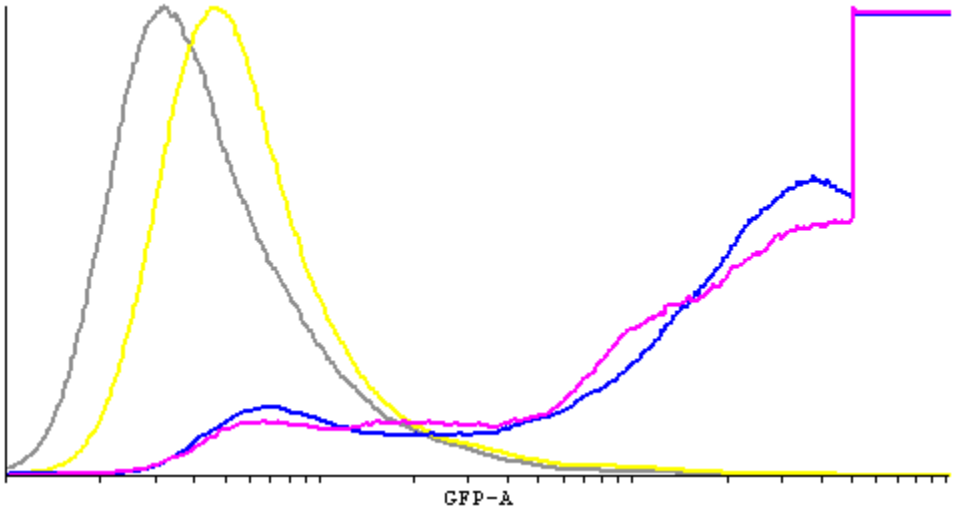


Figure 19: FACS Analysis of PL-GFP and CMV-GFP.
This graph represents the negative control (PL-GFP) and positive control (CMV-GFP). The yellow line represents cell with no LMP1 induction (+Tet), the grey line represents cells with LMP1-induction (-Tet), the purple line represents CMV-GFP with no LMP1 induction (+Tet), and blue indicates LMP1 induction (-Tet).

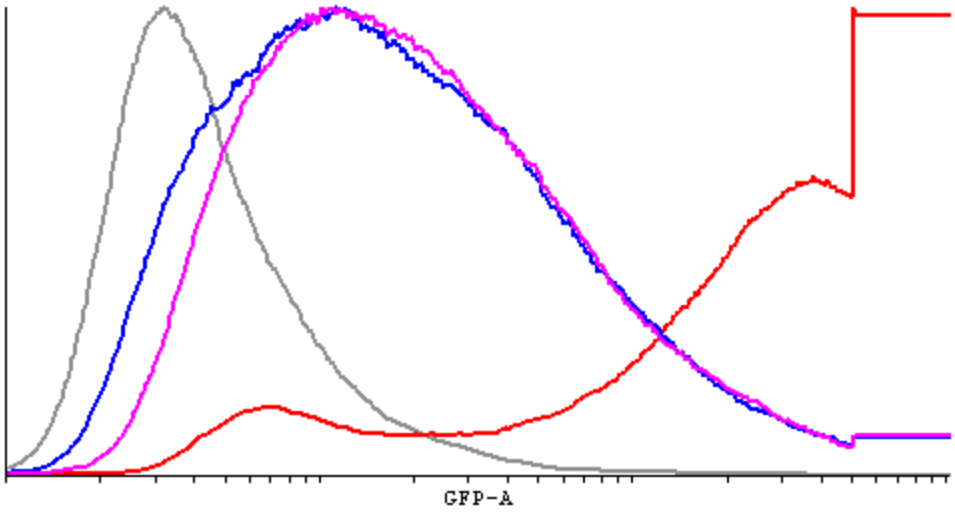


Figure 20: FACS analysis of negative and positive controls versus p2566-GFP.
This graph represents the negative control (PL-GFP) in grey, the positive control (CMV-GFP) in red, p2566-GFP transfected cells with no LMP1 induction (+Tet) in blue, and p2566-GFP transfected cells with LMP1-induction (-Tet) in purple.

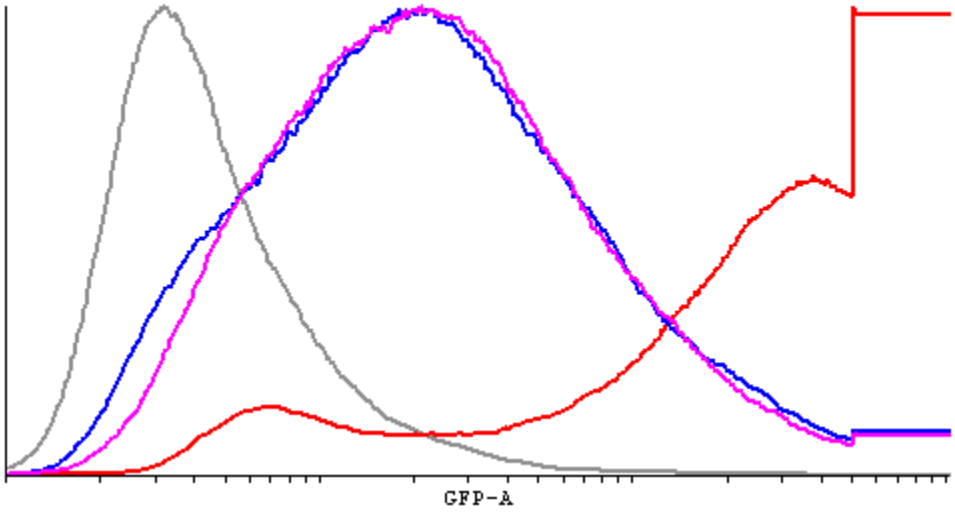


Figure 21: FACS analysis of negative and positive controls versus p1307-GFP.
This graph represents the negative control (PL-GFP) in grey, the positive control (CMV-GFP) in red, p1307-GFP transfected cells with no LMP1 induction (+Tet) in blue, and p1307-GFP transfected cells with LMP1 induction (-Tet) in purple.

RT-qPCR analysis of GFP mRNA

As seen in the above Western blots, changes in protein expression demonstrate a clear TRF2 promoter-GFP downregulation when LMP1 is expressed. Consequently, analysis of GFP

RNA levels was also performed, as it was shown in Lajoie et al. (2015) that LMP1 downregulates both protein and RNA levels of TRF2 and other shelterin proteins. In order to perform this analysis, total RNA was extracted from the cell culture transfected with p1307-GFP, p2566-GFP, negative control pPL-GFP and positive control pCMV-GFP. Using primers GFP F2-R2 for qPCR analysis, the expression levels of GFP RNA were measured (Figure 22). Problems arose during RT-qPCR, in which the primers weren't specific enough to detect the change in RNA expression levels and therefore, we could not accurately distinguish the readings.

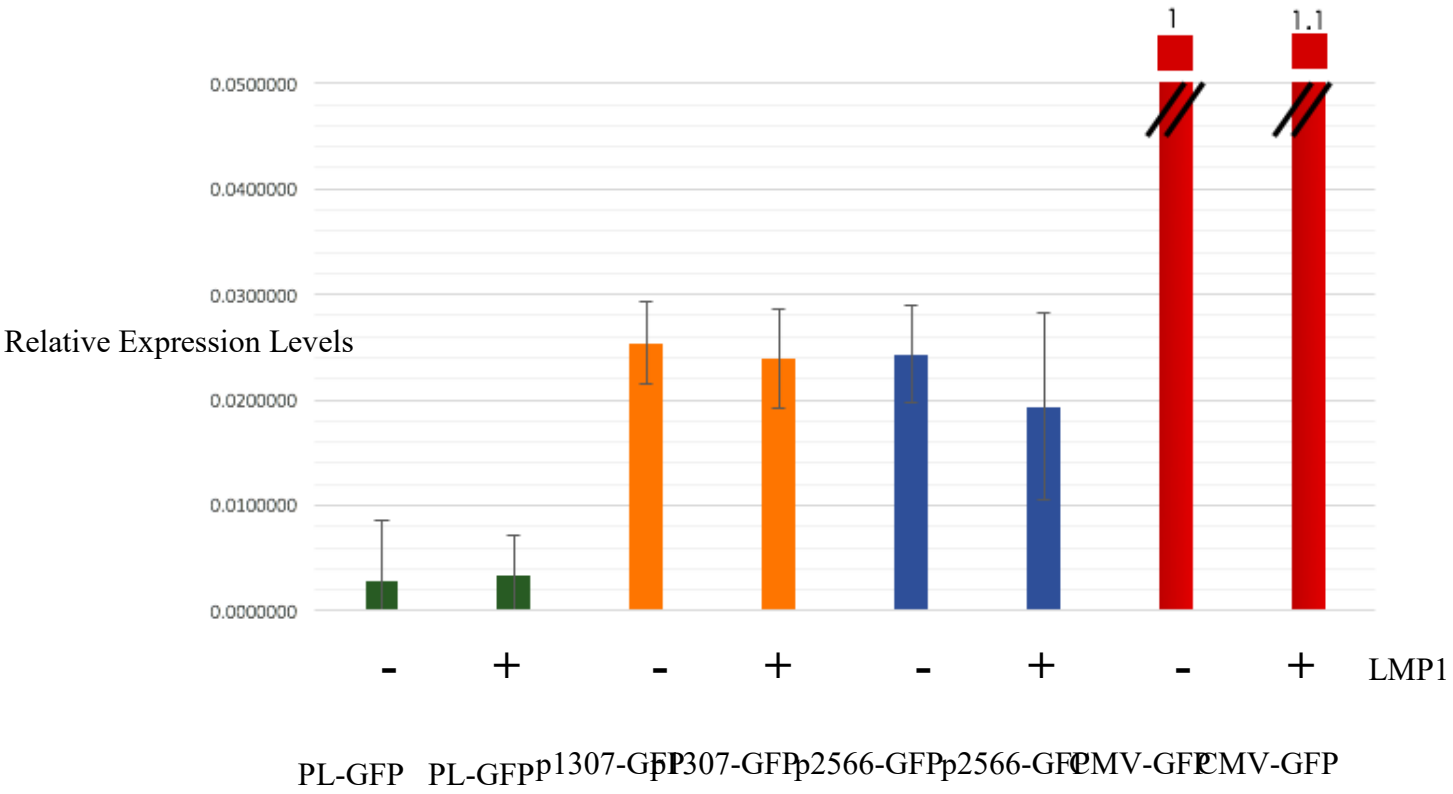


Figure 22: RT-qPCR results of quantification of mRNA GFP in the BJAB tTA LMP1 cells transfected with the different promoter constructs.
One example of RT-qPCR analysis between the PL-GFP, p1307-GFP, p2566-GFP and pCMV-GFP using qPCR GFP F2-R2 primers. Error bars indicate the reproducibility of three independent experiments.

RNA levels for TRF2 and LMP1 were also measured via RT-qPCR to determine if the experiment could be reproducible from previous work that clearly showed downregulation of the endogenous TRF2 and the upregulation of LMP1 when induced. In Figure 23, we see that the

endogenous TRF2 is downregulated upon LMP1 induction by the removal of tetracycline, as shown in Figure 24. In each experiment where a plasmid was transfected for GFP analysis, a downregulation of endogenous TRF2 is observed in the presence of LMP1-induction. The levels of TRF2 differ depending on the expression of LMP1. Since no apparent change was observed using quantitative PCR for mRNA GFP levels, another method known as ddPCR was used.

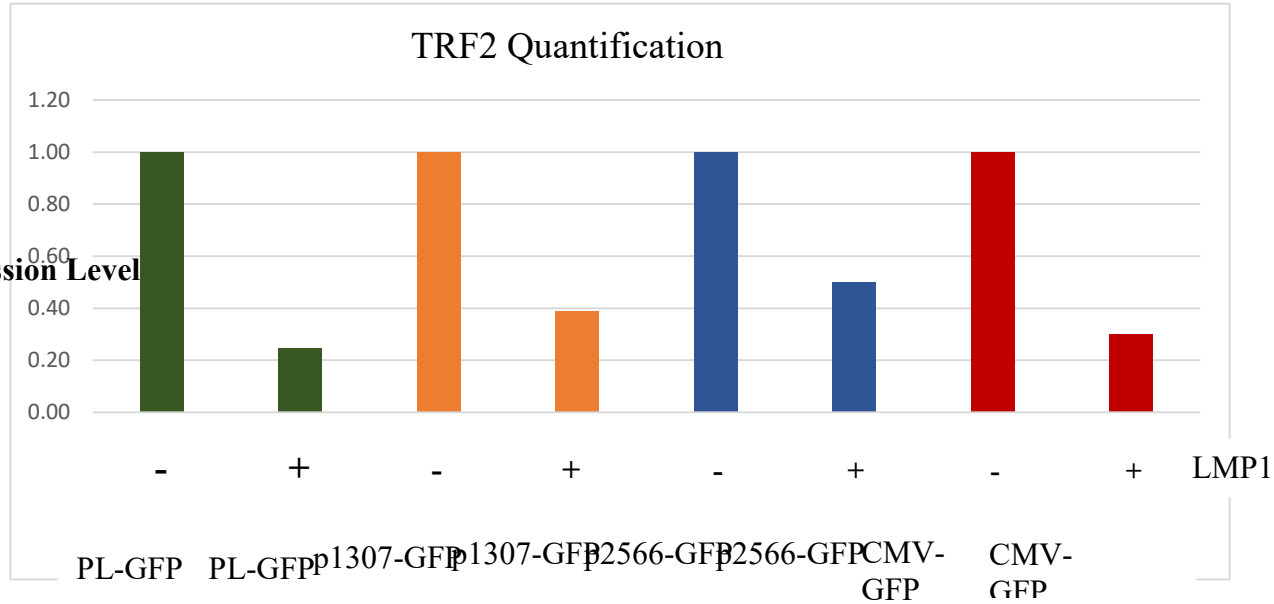


Figure 23: RT-qPCR quantification of mRNA TRF2.
An example of TRF2 levels in BJAB tTA cells transfected with the different promoter constructs. After transfection with promoter PL-GFP, p1307-GFP, p2566-GFP and pCMV-GFP, and RNA extraction was performed and by RT-qPCR, TRF2 levels were measured using primers called TRF2 For and TRF2 Rev.

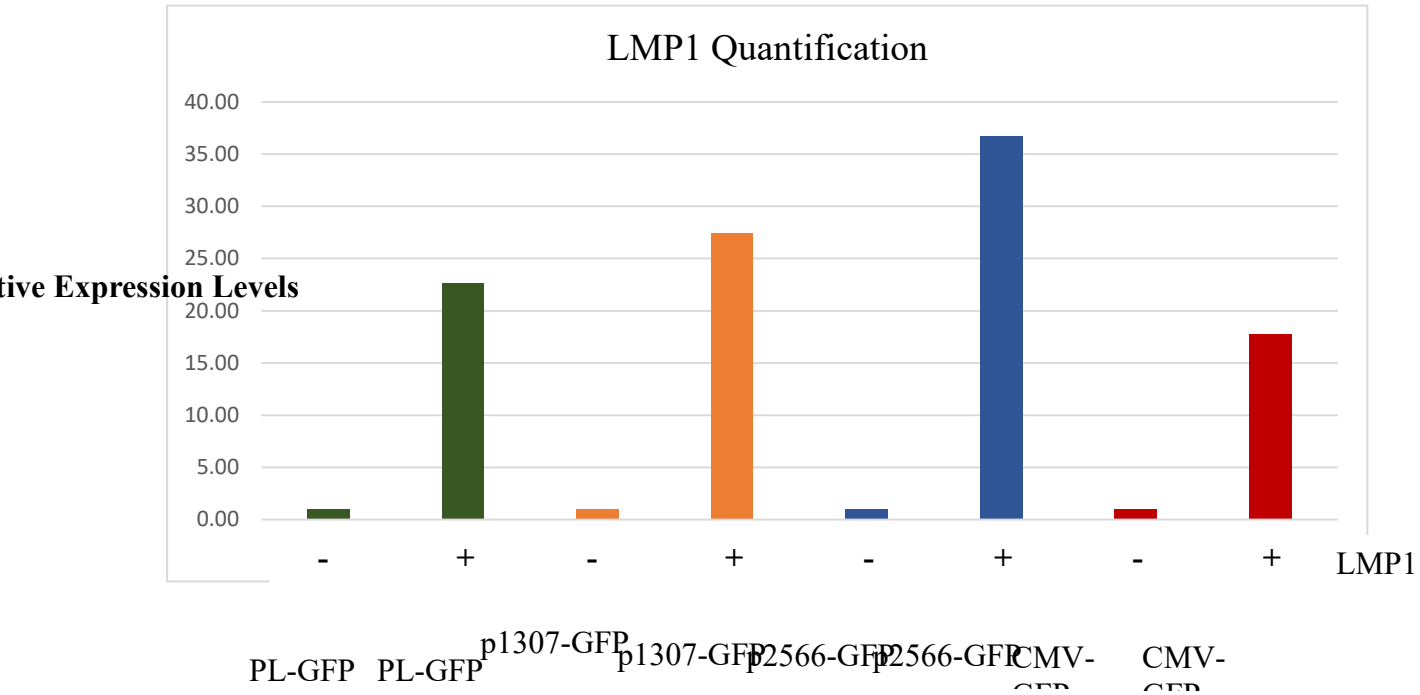


Figure 24: RT-qPCR quantification of mRNA LMP1.
An example of LMP1 levels in BJAB tTA LMP1 transfected with the different promoter constructs. After transfection with promoter PL-GFP, p1307-GFP, p2566-GFP and pCMV-GFP, and RNA extraction was performed and by RT-qPCR, LMP1 levels were measured using primers called LMP1 For and LMP1 Rev.The relative mRNA expression levels are represented.

GFP mRNA analysis through ddPCR

Droplet digital PCR (ddPCR) is a method in which the cDNA sample is fractionated into water-oil droplets and in each individual droplet, there is PCR amplification of the targeted template molecule that can be measured (Taylor, Laperriere & Germain, 2017). Figure 25-28 are the results from the transfection of PL-GFP, p50-GFP, p100-GFP, p115-GFP, p125-GFP, p140-GFP and p150-GFP in which GFP, TRF2 and LMP1 levels were measured using ddPCR. Primers GFP F16-R16 were used as a first trial. As seen in Figure 25, no downregulation between no LMP1 and with LMP1 samples was observed. The assay was performed with another set of primers, GFP F18-R18, as seen in Figure 26. Again, there was no observable downregulation between non-induced and induced LMP1 samples. However, LMP1 detection with ddPCR in Figure 27 showed clear upregulation in samples containing LMP1 induction. Endogenous TRF2 levels were also measured as an internal control for the experiment. TRF2

showed clear downregulation between non-induced versus induced LMP1 (Figure 28). These data therefore strongly suggest that the primers used in ddPCR weren't specific enough for our GFP analysis. This was determined as RNA levels in the plasmid containing no promoter upstream the GFP gene, had expression level comparable to cell containing a promoter in front of the GFP gene. No further experiments were performed with ddPCR or qPCR.

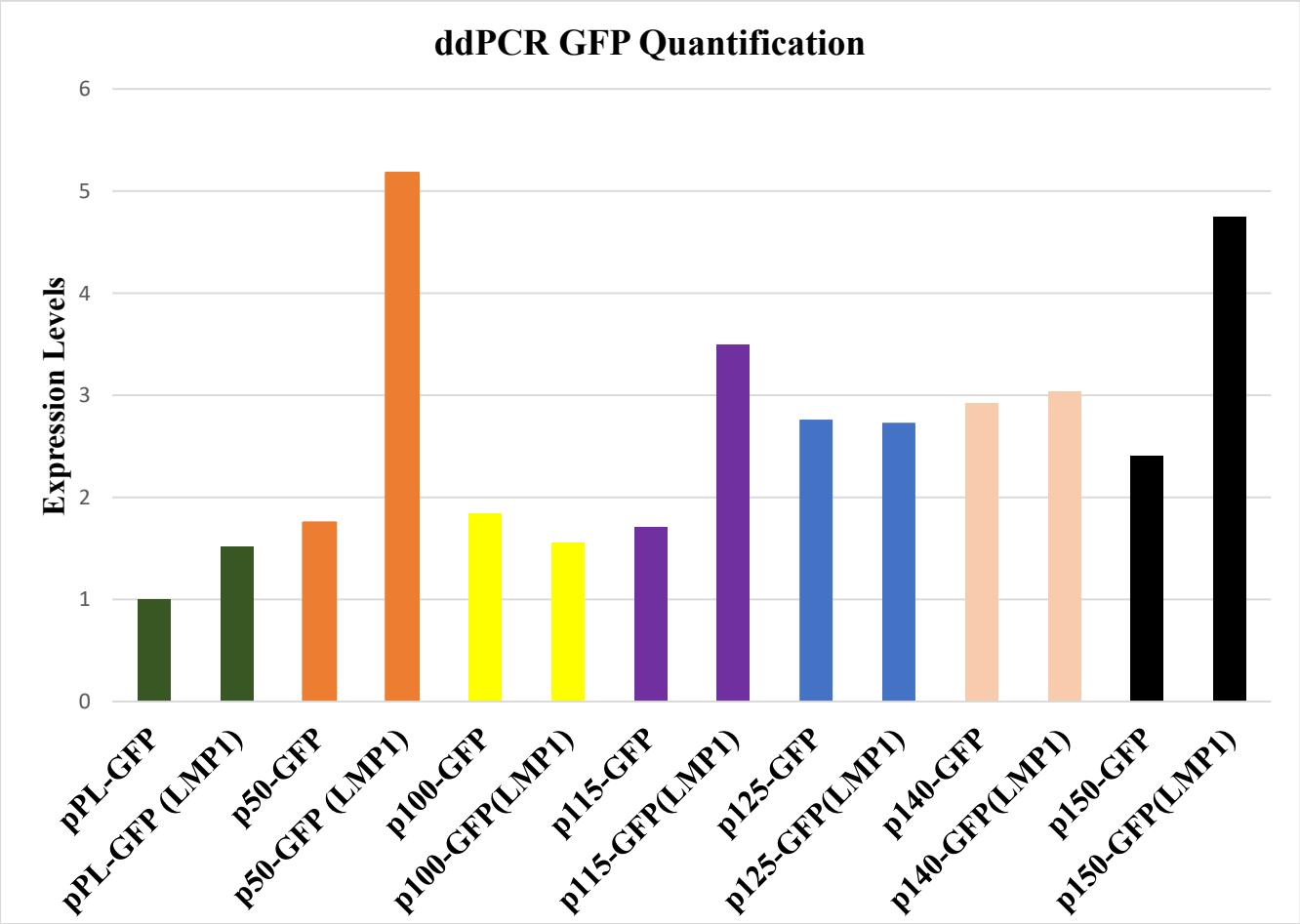


Figure 25: ddPCR analysis of expression levels of GFP RNA using primers GFP F16-R16
The number on the side represents the plasmid transfected into the BJAB tTA LMP1 cell line. The sign beside each name represents the sample induced with and without LMP1 during the experiment.

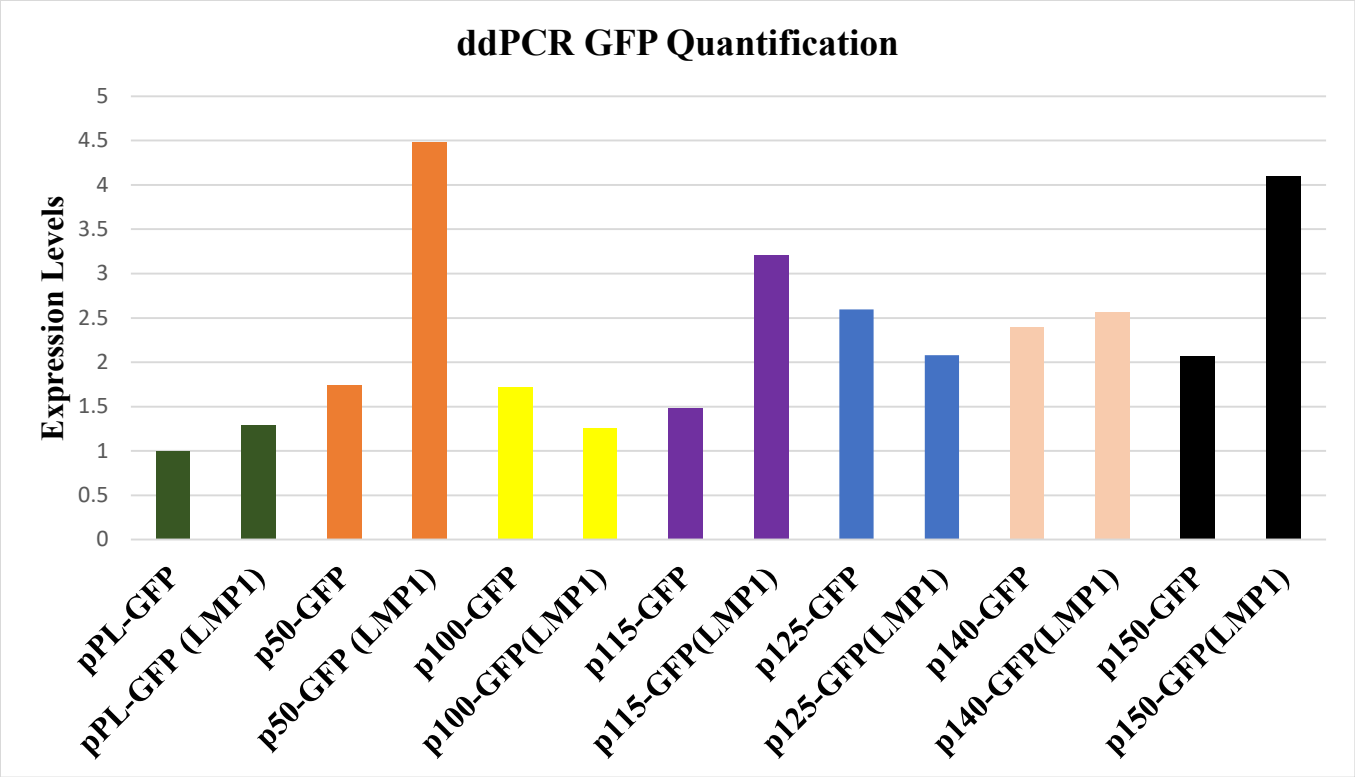


Figure 26: ddPCR analysis of expression levels of GFP RNA using primers GFP F18-R18
The number on the side represents the plasmid transfected into the BJAB tTA LMP1 cell line. The sign beside each name represents the sample induced with and without LMP1 during the experiment.

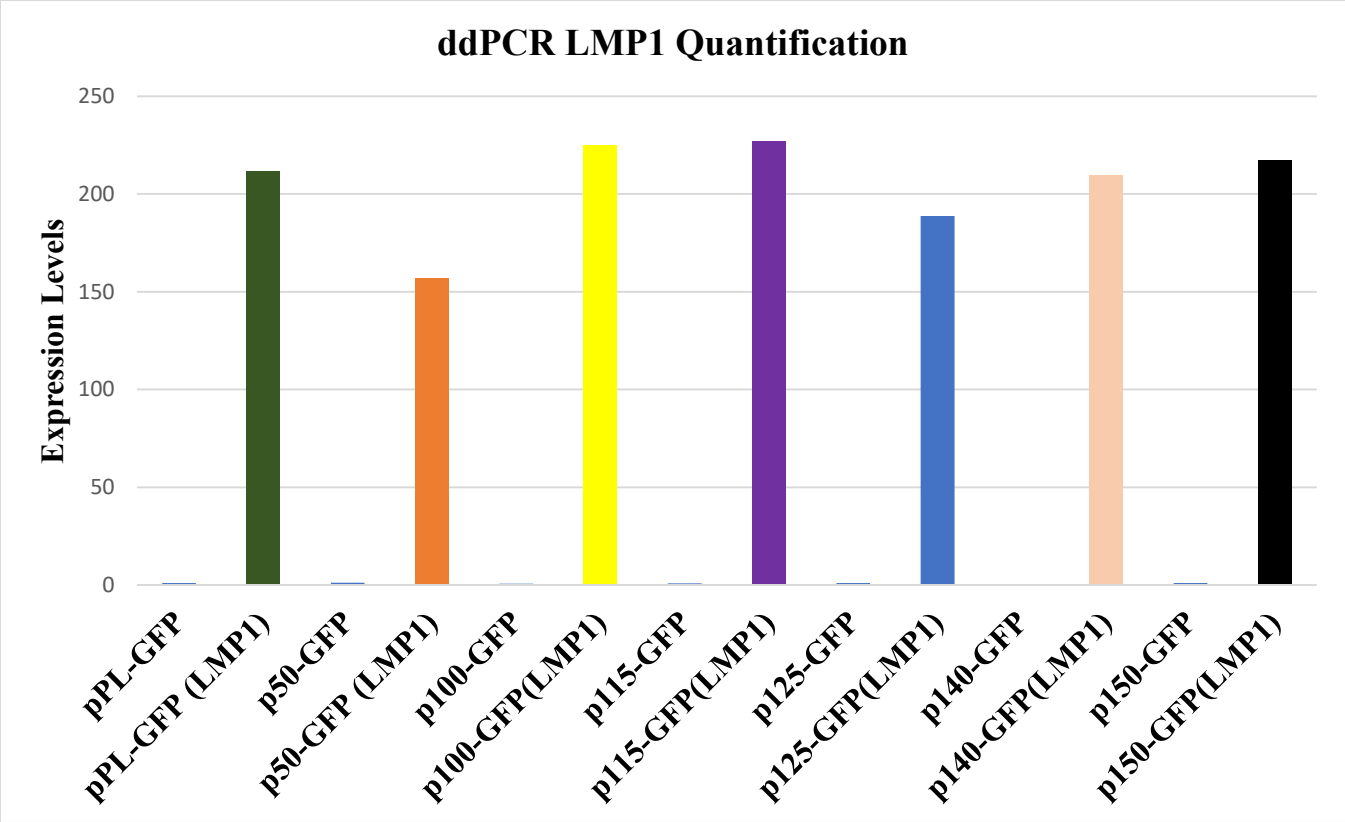


Figure 27: ddPCR analysis of expression levels of LMP1 RNA
The number on the side represents the plasmid transfected into the BJAB tTA LMP1 cell line. The sign beside each name represents the sample induced induced with and without LMP1 during the experiment.

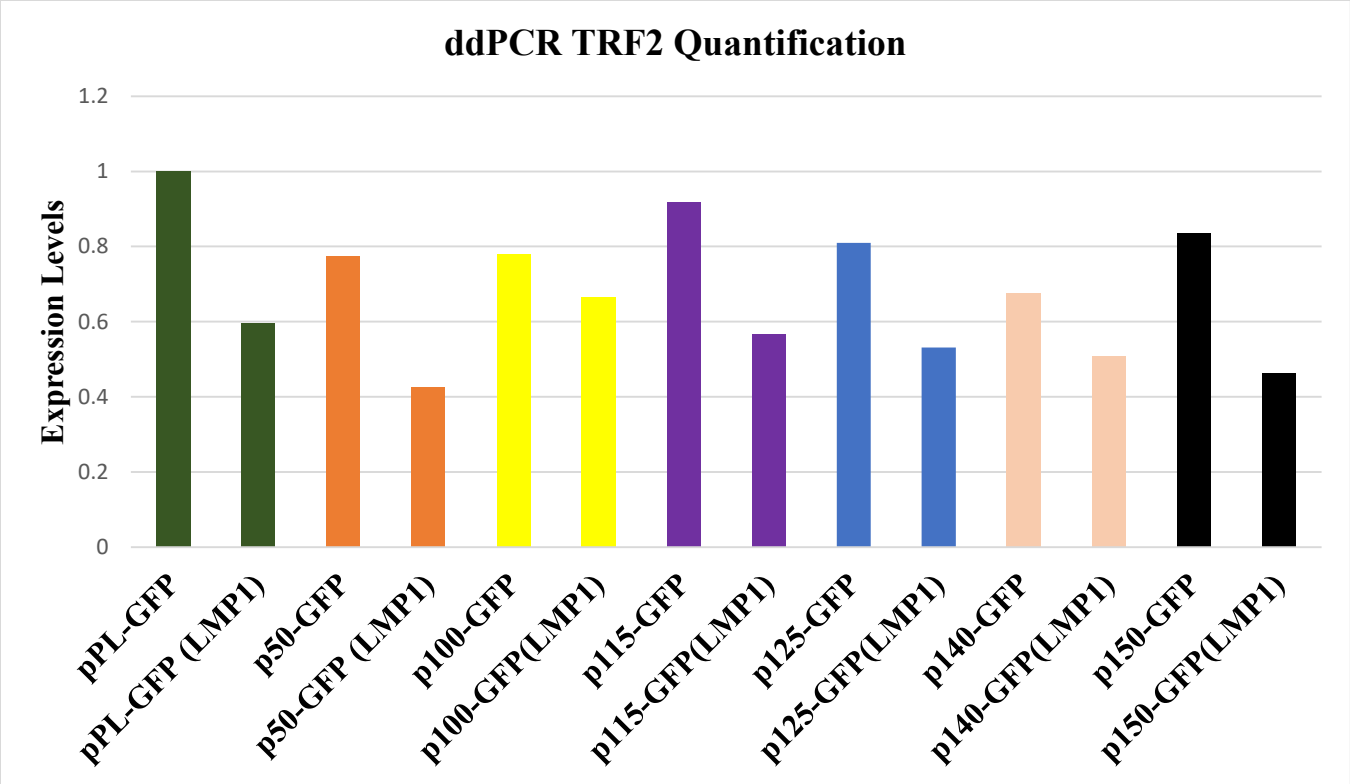


Figure 28: ddPCR analysis of expression levels of TRF2 RNA
The number on the side represents the plasmid transfected into the BJAB tTA LMP1 cell line. The sign beside each name represents the sample induced with and without LMP1 during the experiment.

Results-Chapter 2

Localization and analysis of a smaller region of the TRF2 promoter regulated by LMP1-induction

Preamble

The first objective of this project was answered by the Western blot shown in Figure 17. That result demonstrated that LMP1 induction in the BJAB tTA LMP1 cell line caused LMP1 to repress the TRF2 promoter. Once this was determined, the next objective was to localize a smaller region of the TRF2 promoter where LMP1 indirectly affects TRF2 at the promoter. To answer this question, Western blots for GFP protein analysis were used to determine if a region in which LMP1 did not affect the regulation any longer could be delineated. Once the region was localized, EMSA, DNA affinity chromatography and mass spectrometry were used for protein analysis in the identified significant region in an attempt to find regulatory factors.

Identification of a promoter region affected by LMP1-induction

To proceed with the analysis, constructs with various extents of the TRF2 gene promoter ranging from 821 bp down to 50 bp were made to identify where LMP1 induction stops regulating the TRF2 promoter. It was hypothesized that LMP1-induction might be interfering with an enhancer or a region of regulation. The first construct was made by using plasmid p1307-GFP, in which the plasmid was digested with restriction enzyme XhoI and EcoRV, to remove a 486 bp fragment from the region between 1307bp and 822bp. Using a Klenow treatment, 3'overhangs were removed or 3' resected (5' overhang) ends were filled in to create blunt ends. The vector was then ligated to form a circularized plasmid and sent for sequencing. This created the plasmid called p821-GFP. Plasmid p737-GFP was made by PCR amplification by using primers F24 and R13 and as the DNA template, the plasmid p821-GFP was used. The same plasmid, p821-GFP, was then used as a vector by linearizing with restriction enzymes XhoI and HindIII, to then insert the 737 bp promoter fragment. The same procedure was done for plasmid p614-GFP except primers F25 and R13 were used, plasmid p680-GFP primers F27 and R13 was used and for plasmid p484-GFP, primers F28 and R13 were used. All plasmids were sent for sequencing before use in downstream applications. The BJAB tTa cells were transfected (nucleofection) using the program EN-150. After transfection, the sample was split in two. Half

went into a media with tetracycline and the other half into tetracycline-free media. A 10% SDS-PAGE gel was prepared for Western blot analysis. The membrane was probed with anti-LMP1, anti-GFP, anti-actin and anti-tubulin as housekeeping gene loading controls (Figure 29).

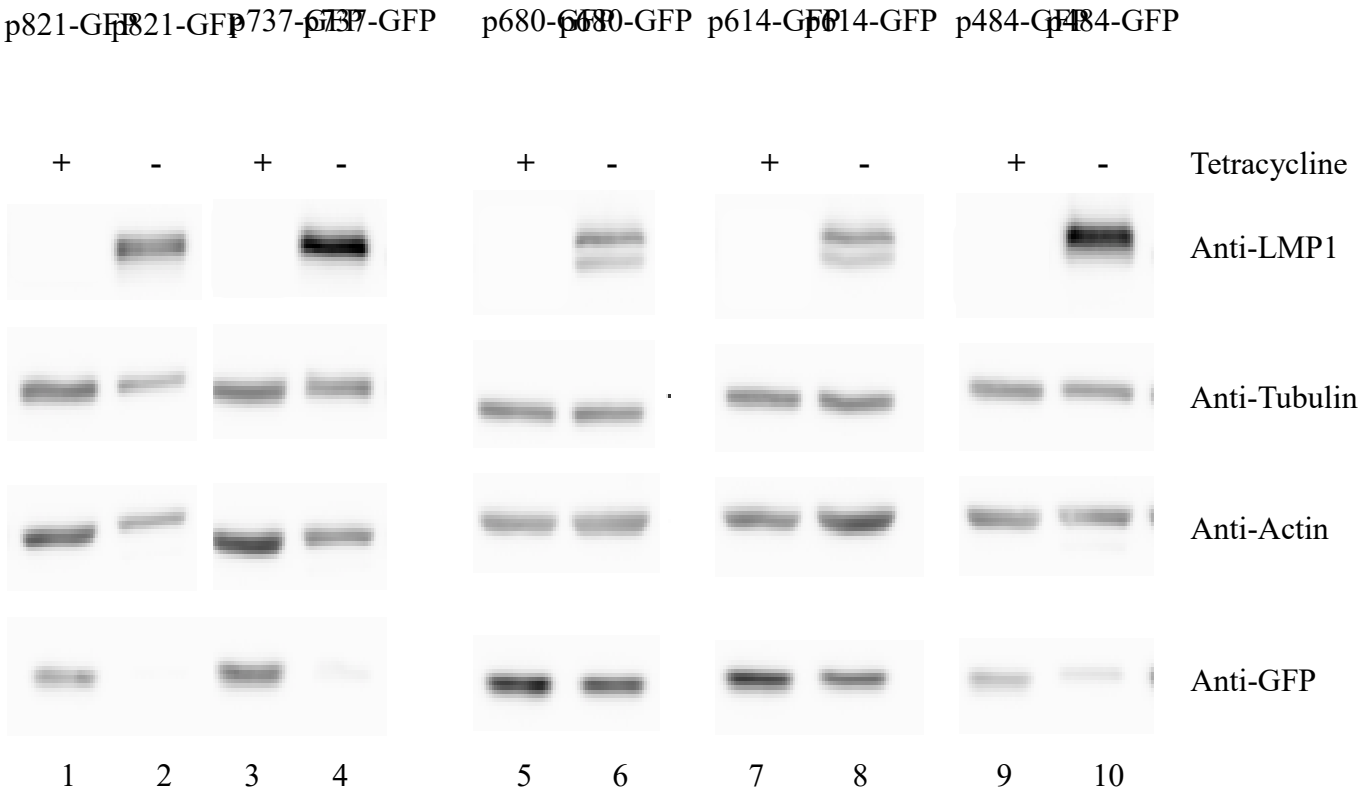


Figure 29: Western Blot of BJAB tTA LMP1 transfected with TRF2 promoter plasmids containing different promoter lengths.
Transfection with the different plasmid constructs of the TRF2 promoter. Membrane was probed with anti-LMP1, anti-GFP, anti-actin and anti-tubulin. Lanes 1-2 were cells transfected with a TRF2 promoter of 821bp, lanes 3-4 were cells transfected with a TRF2 promoter of 737bp, lanes 5-6 were cells transfected with a TRF2 promoter of 680bp, lanes 7-8 were cells transfected with a TRF2 promoter of 614bp, lanes 9-10 were cells transfected with TRF2 promoter of 484bp.

Downregulation of GFP was still observed with promoter constructs of 821bp, 727bp, 680bp, 614bp and 484bp. There is weak downregulation from promoter of length 680bp and 614 bp since the experiments depend on the strength in which LMP1 is induced. The more LMP1-induction, the more it will exert its effect of the TRF2 promoter. A double band can also be

observed in the anti-LMP1. This band is from anti-Tubulin. LMP1 and tubulin migrate at a height of 65kDa and 70kDa which are pretty close together. If the antibodies are used on the same western blot, and depending on the migration and separation, a double band can be seen. We can also see a stronger signal for LMP1 in lanes 4 and 10 because Tubulin and LMP1 are at the same height. Nevertheless, the construct p484-GFP, shows downregulation in three different experiments. Therefore, it was decided to further dissect the promoter in order to localize a region in which LMP1-induction no longer affected the expression of GFP protein via the TRF2 promoter. Using the p821-GFP as DNA template with forward and reverse primers, whole plasmid PCR was performed in order to obtain a long fragment that can be ligated to form a circularized plasmid. Plasmids p400-GFP, p300-GFP, p250-GFP and p200-GFP were produced using their respective forward primers as mentioned in table 7 and with the reverse primer ALL-RW R3. Figure 30 represents the Western blot results after transfection of the plasmids into BJAB tTA cell line induced (-Tet) and non-induced (+Tet) with LMP1. The membrane was probed with anti-LMP1 and anti-GFP, as well as anti-actin and anti-tubulin for housekeeping gene loading controls.

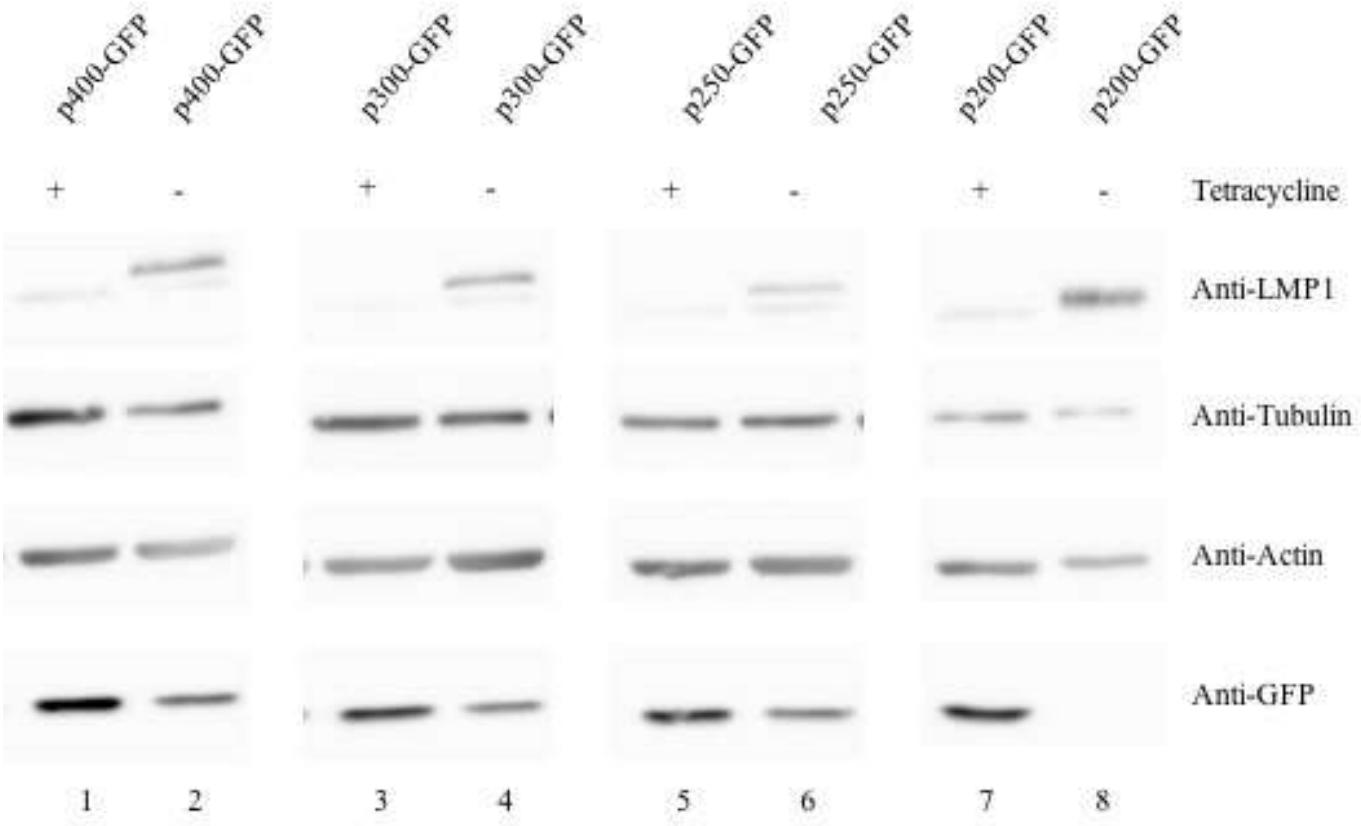


Figure 30: Western Blot of BJAB tTA LMP1 transfected with TRF2 promoter plasmid containing different promoter lengths.
Transfection with the different plasmid constructs of the TRF2 promoter. The membrane was probed with anti-LMP1, anti-GFP, anti-actin and anti-tubulin. Lanes 1-2 were cells transfected with TRF2 promoter 400bp, lanes 3-4 were cells transfected with TRF2 promoter 300bp, lanes 5-6 were cells transfected with TRF2 promoter 250bp and lanes 7-8 were cells transfected with TRF2 promoter 200bp.

At 150bp of the promoter region, a downregulation is observed when LMP1 is induced (Figure 31). When the promoter regions of 100bp and 50 bp were examined, basal GFP transcription similar to background levels was observed (Figure 31). Depending if our GFP primary antibody was fresh or older, the levels of GFP expression for 50bp would vary. The GFP expression would be visible on the blot with a fresh GFP, but if extra time for washing were added after incubation of the secondary antibody, the GFP level were not detectable on the blot. Nevertheless, the GFP expression was weak or undetectable on both p100-GFP and p50-GFP when the experiment was repeated multiple times. It was decided to dissect in between 100 and 150bp to see if a more precise region could be localized for further analysis.

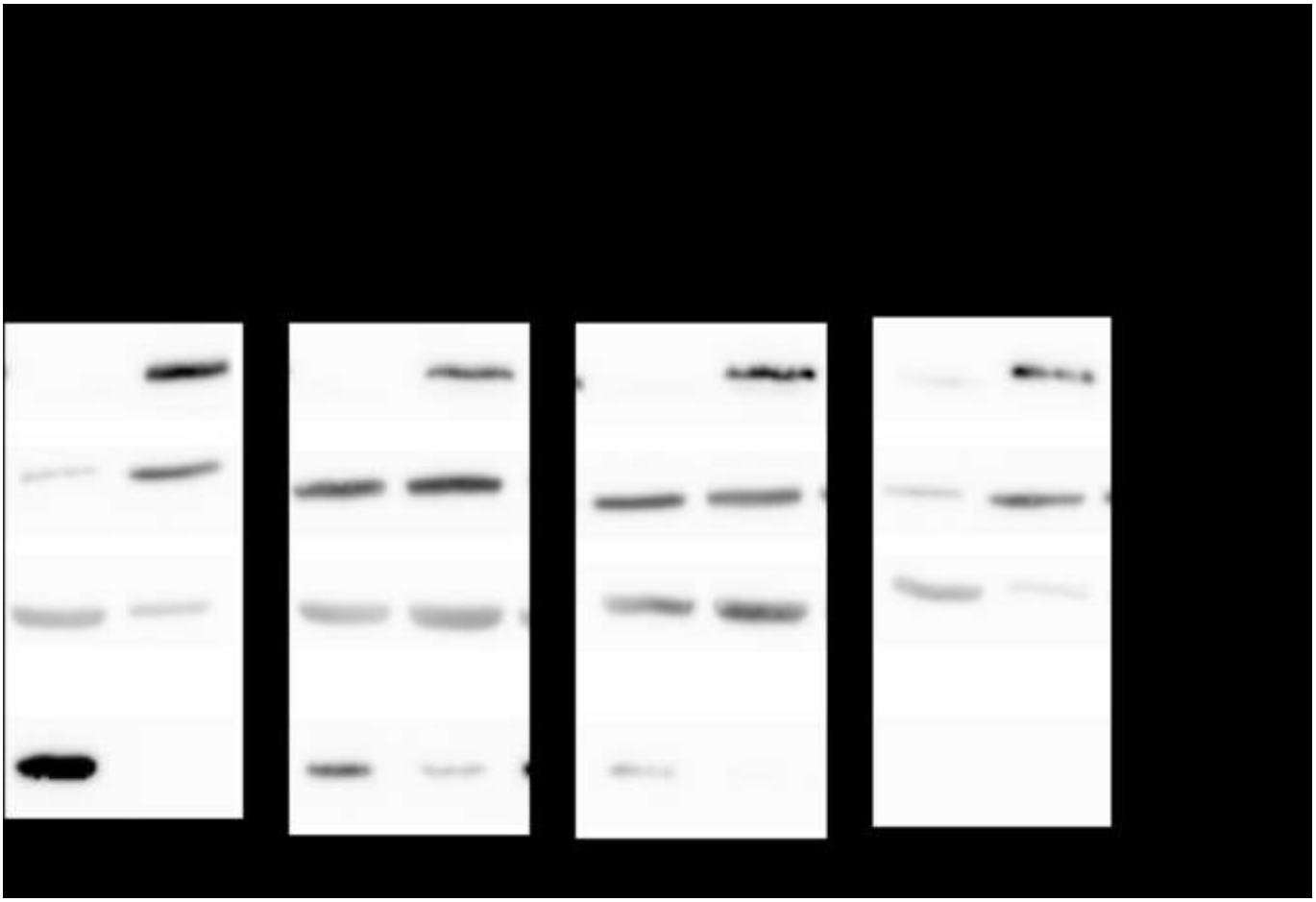


Figure 31: Western Blot of BJAB tTA LMP1 transfected with TRF2 promoter plasmid containing different promoter lengths.
Transfection with the different plasmid constructs of the TRF2 promoter. Membrane was probed with anti-LMP1, anti-GFP, anti-actin and anti-tubulin. Lanes 1-2 were cells transfected with TRF2 promoter 150bp, lanes 3-4 were cells transfected with TRF2 promoter 100bp, lanes 5-6 were cells transfected with TRF2 promoter 50bp and lanes 7-8 were cells transfected with no promoter in front of the GFP gene.

Using whole plasmid PCR as mentioned in the materials and methods, constructs of 115bp, 125bp and 140bp promoters were made. The results shown in Figure 32 are representative of three independent experiments for which one representative Western blot is shown. In Figure 32 lanes 13-14 (150 bp promoter) a downregulation was observed upon LMP1 induction. In lanes 11-12, the 140bp promoter also shows downregulation in the presence of LMP1, which is again seen for the 125bp promoter in lanes 9-10. Significantly, low basal transcription of GFP is seen for the 100 bp promoter (lanes 5-6) and 115 bp promoter (lanes 7-8). The experiment was repeated three times, each time showing similar results. From this, it was determined that a 125

bp promoter length is needed for sufficient GFP transcription. When the promoter section is below 125 bp, minimal basal transcription leads to GFP levels undetectable under the Western blot conditions used.

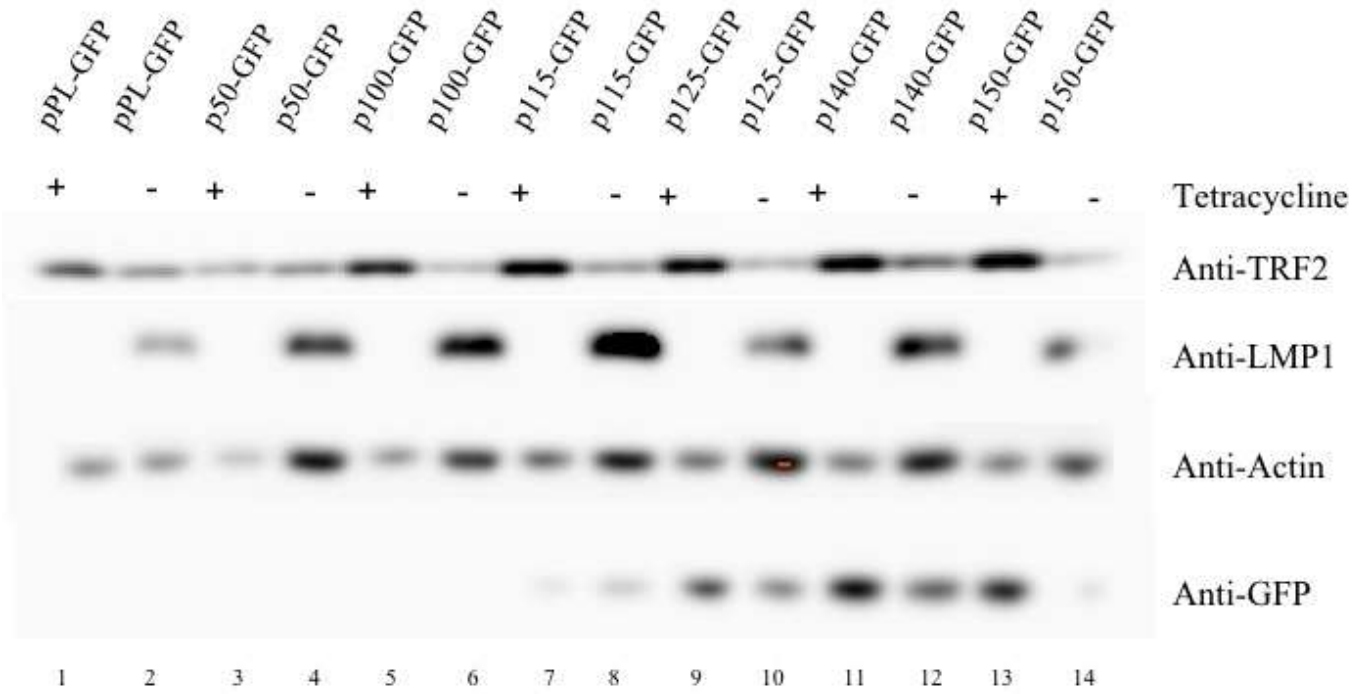


Figure 32: Western Blot of BJAB tTA LMP1 transfected with TRF2 promoter plasmid containing different promoter lengths.
Transfection with the different plasmid constructs of the TRF2 promoter. Membrane was probed with anti-LMP1, anti-GFP, anti-actin and anti-tubulin. Lanes 1-2 are cells transfected with no promoter in front of the GFP gene, lanes 3-4 are cells transfected with 50 bp of TRF2 promoter, lanes 5-6 are cells transfected with 100 bp of TRF2 promoter, lanes 7-8 were transfected with 115 bp of TRF2 promoter, lanes 9-10 were transfected with 125 bp of TRF2 promoter, lanes 11-12 were transfected with 140 bp of TRF2 promoter, lanes 13-14 were transfected with a TRF2 promoter of 150 bp.

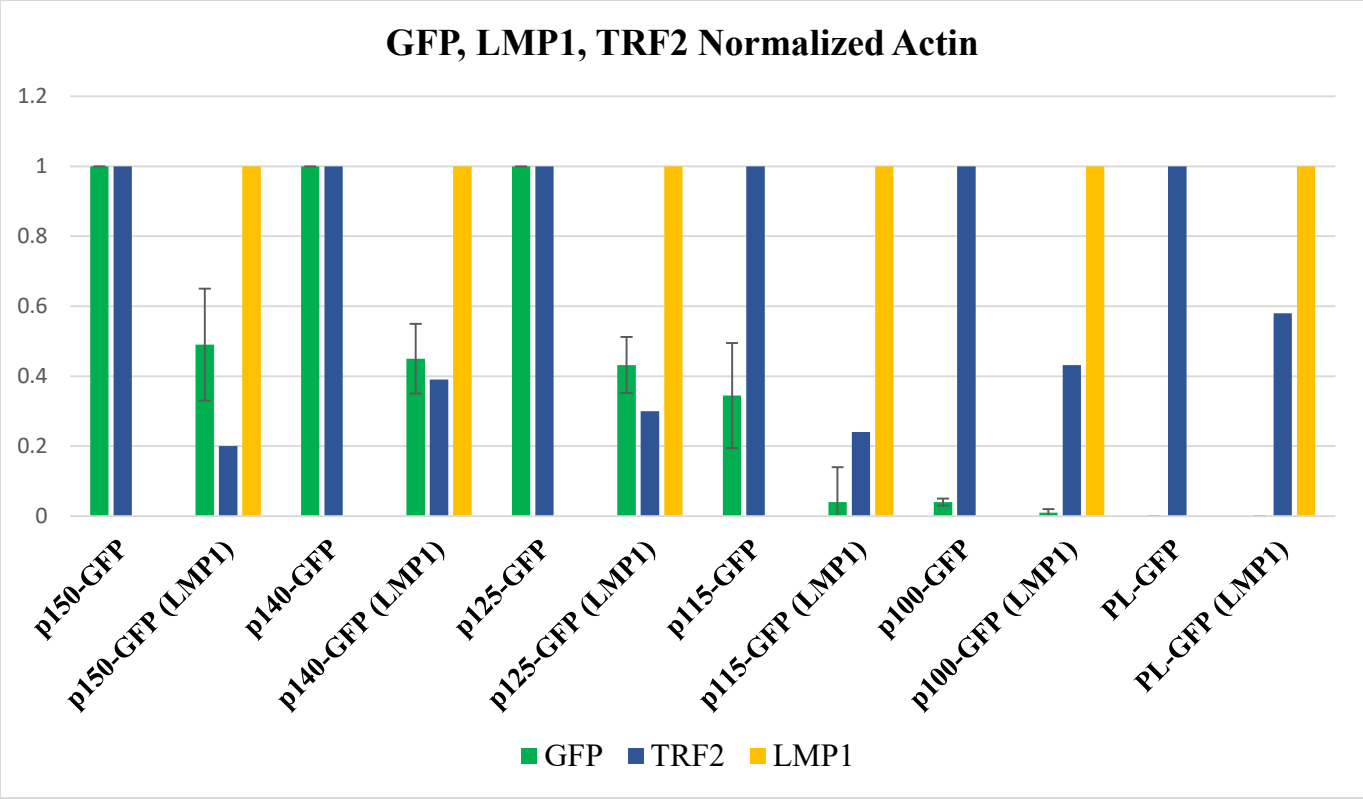


Figure 33: Quantification of three independent western blots for GFP analysis.
The green line represents the GFP expression of three independent experiments quantified (using image.j) and normalized against actin. The blue lines represent the expression levels of TRF2 of one experiment from western blot in Figure 29. The yellow lines represent the expression levels of LMP1-induced in the cell line.

The TRF2 promoter was dissected from a region of 2566 bp down to 50 bp where downregulation could be observed down to 125 bp. If the promoter is shorter than 125 bp, low basal GFP transcription is seen. A quantification of three independent experiments for GPF analysis was performed (Figure 33). The goal of our experiments was to delineate a region in which LMP1-induction did not affect the regulation of TRF2 at its promoter. In this case, no matter how low we dissected the promoter, downregulation could be observed. The promoter at length 115bp, was also at limit for GFP detection and showed downregulation. Thus, showing that for transcription of GFP gene, at least 125bp is needed for the basal transcription machinery to be able to attach and transcribe. At 100bp and 50bp of promoter length, basal background GFP expression was undetectable. What we were looking for was a region not affected by LMP1-induction but instead, no matter how short the promoter length was, LMP1 was having an effect. Therefore, we hypothesize that LMP1-induction could interfere at the beginning of the promoter

either at the level of the basal transcription machinery or with a transcription regulator. To further study the promoter, it was decided to take a region between -130 bp and -100 bp of the TRF2 promoter. This region provides 30 bp of DNA for analysis by EMSA to first confirm if protein complexes bind to the region. Secondly, upon confirmation of protein binding, DNA Affinity Chromatography analysis of nuclear extract could be performed to provide further information on the protein complexes bound when LMP1 is present and absent.

Electrophoretic-mobility shift assay (EMSA)

EMSA is a technique used to study protein-DNA interactions. By using a 5' γ-ATP^{32} radiolabelled dsDNA probe, band patterns can be observed and compared to nuclear extract from cells with LMP1 induced and non-induced. A region of 30 bp was chosen between -130 and -100 bp of the TRF2 promoter to study the effects of LMP1 induction in the cell. An EMSA was performed to first determine if protein complexes bind to the region previously determined to be significant by Western blot analysis of GFP levels. As per description of the procedure found in the materials and methods, an EMSA was performed (Figure 34). Lane 2 contains nuclear extract without LMP1 induction, meaning the BJAB tTA LMP1 cells were incubated in media with tetracycline. Lanes 3-5 contain samples where 10x, 50x or 200x fold excess of the specific competitor was added. The yellow stars in Figure 34 indicate the bands that are disappearing when excess of a specific competitor was added, in this case, it was unlabelled 30 bp dsDNA. A non-specific competitor was also added in excess fold of 50x or 200x to determine if the same bands that are disappearing are disappearing with the specific competitor will still be present when a non-specific competitor is added. As can be observed in lanes 7-8, the bands that disappear in lanes 4-5 are still present in lanes 7-8. This confirms that the protein complexes found in lanes 3-5 are most likely complexes specific to the 30 bp dsDNA probe. Lane 9 contains nuclear extract with LMP1 induced, meaning the BJAB tTA LMP1 cells were incubated in media without tetracycline. Lanes 10-12 contain samples where 10x, 50x or 200x fold excess of the specific competitor was added. The white stars in Figure 34, indicate the bands that disappear when an excess of a specific competitor is added. A non-specific competitor was also added in excesses of 50x or 200x fold excess. As can be observed, the bands that disappear in lanes 10-12

are still present in lanes 14-15. This confirms that the protein complexes found in lanes 3-5 are specific complexes binding to the 30 bp region of the dsDNA probe.

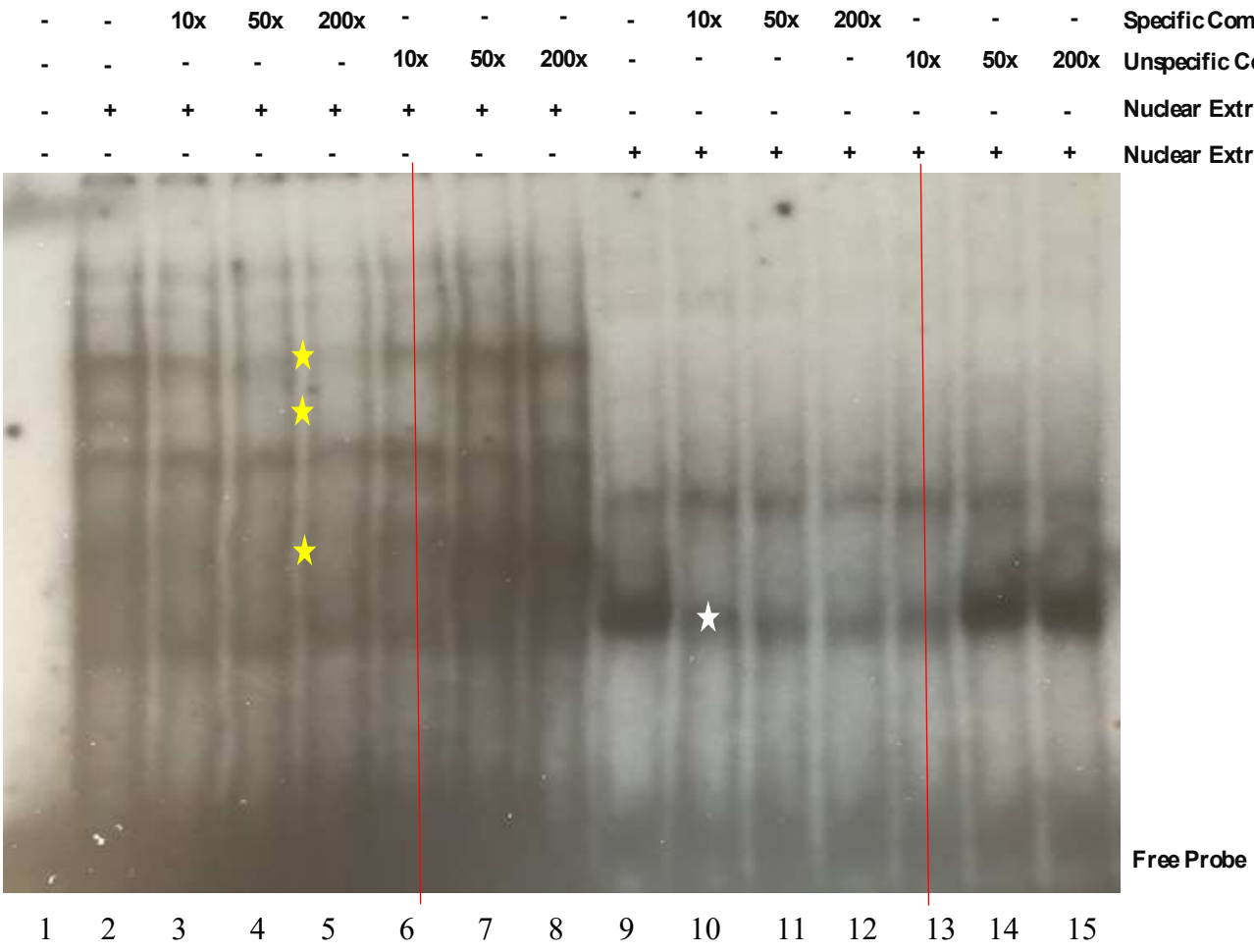


Figure 34: Electrophoretic-mobility shift assay (EMSA) with the 30bp dsDNA
An EMSA was performed using an oligonucleotide probe containing the region between -130 bp and -100 bp of the TRF2 promoter starting from the ATG start site as 1. Lanes 2-8 are with nuclear extract with No LMP1 (+Tet) and lanes 9-15 are with nuclear extract with LMP1 induction (-Tet). Lanes 3-5 and lanes 10-12 used a specific competitor (unlabelled 30 bp dsDNA). Lanes 7-8 and lanes 14-15 used a non-specific competitor (Oligo B). Disregard lanes 6 and 13 from the data analysis as the wrong competitor was added during the experiment.

The EMSA data analysis confirmed that there is a change of proteins in the nuclear extract in the presence or absence of LMP1 induction. Consequently, an RNA sequencing assay (RNA-Seq) was performed to determine what other genes could be downregulated in the presence of LMP1 induction in the BJAB tTA LMP1 cell line.

RNA Sequencing

LMP1 affects multiple pathways in the cell for the purpose of growth, proliferation and survival of the Epstein-Barr virus. The LMP1 oncogenic protein was found to effect multiple pathways such as activating the NF-kB, JNK, and p38 (Soni, Cahir-McFarland & Kieff, 2007). For the purposes of this study, only downregulated genes were focused on. TRF1 and TRF2, important telomeric proteins, were found to be downregulated when LMP1 was induced in the cells. To find a potential pathway affecting these genes, all RNAs extracted were studied to find what other genes were downregulated by LMP1-induction. Table 8 represents a list of the top 26 genes affected by LMP1 induction. RNA seq data confirm the expected TRF1 and TRF2 downregulation upon induction of LMP1. Congruent with previously published results (Lajoie et al., 2015), no significant change was observed in POT1 levels in LMP1-induced and non-induced cells.

Gene	Q-Value	Log2-ratio (No LMP1 versus LMP1 sample)	Reads Sample 1 (No LMP1)	Reads Sample 1 (LMP1)	Reads Sample 2 (No LMP1)	Reads Sample 2 (LMP1)
JCHAIN	2.36E-01	3.9	20.4	1.3	23.2	1.7
SPINK2	2.55E-01	3.3	163.8	14.5	219.1	23.9
LTB.d	2.25E-01	3.2	55.3	5.8	52.3	5.5
AIF1.b	2.40E-01	3.0	11.7	1.7	13.8	1.5
KCNN3	2.50E-01	3.0	9.3	1.2	11.5	1.3
LDLRAP1	2.74E-01	2.9	23.2	1.7	14.8	3.4
AIF1.d	2.40E-01	2.8	11.8	1.9	13.8	1.9
AIF1.c	2.40E-01	2.8	11.8	1.9	13.8	1.9
AIF1	2.40E-01	2.8	11.8	1.9	13.8	1.9
CA11	2.47E-01	2.7	9.8	1.5	12.0	1.8
ITK	2.25E-01	2.6	11.7	1.9	10.9	1.9
LRMP	2.53E-01	2.6	157.3	28.9	198.2	30.7
DBN1	2.40E-01	2.5	39.2	7.5	45.8	7.5
TRAF3IP3	2.07E-01	2.4	12.1	2.1	12.3	2.5
MED20	3.73E-01	2.4	5.3	2.7	26.3	3.5
PRKCE	3.73E-01	2.3	44.7	8.4	11.4	3.0
RGS13	2.11E-01	2.2	391.3	76.0	369.6	85.4
TNS3	2.25E-01	2.1	17.1	5.6	18.9	2.6
NCAM2	2.77E-01	2.1	9.4	4.8	15.9	1.0
RETREG1	2.25E-01	2.1	108.0	24.5	116.9	28.3
MEIKIN	3.09E-01	2.1	15.7	4.2	29.1	6.4
TRIB1	2.36E-01	2.1	13.5	3.2	15.4	3.8
CD69	2.36E-01	2.1	79.5	14.5	66.3	20.7
BCL6	2.58E-01	2.0	12.4	2.9	16.7	4.2
TERF2	3.73E-01	1.1	13.7	9.7	30.3	10.8
TERF1	2.25E-01	0.6	11.3	7.8	12.0	7.5
POT1	4.64E-01	0.0	37.7	37.4	28.2	22.6

Table 8: RNA-Seq data of the most downregulated genes in LMP1-induced BJAB tTA LMP1 cells.

Data analysis of the RNA-Seq results confirmed that 119 genes had equal to or greater downregulation than TRF2. With this data, bioinformatics analysis was performed to determine if transcription factor common motifs could be found within their promoters. With the help of the Université de Sherbrooke RNomics platform, bioinformatic analysis was carried out between -130 bp and -100 bp of the TRF2 promoter to determine what motifs occur within this region. The JASPAR program was used to perform the analysis. Once found, the motifs were then compared within a region limited to 500bp of each gene promoter to see if some common motifs could be

observed in the downregulated genes. 14 motifs were found within the region between -130 bp and -100 bp of the TRF2 promoter. Table 9 contains the motifs with the total number of occurrences. Overall, 121 genes were analysed as TPP1 and POT1 were added to the list. A value of 1 was given to each gene in which a motif could be found and a 0 value was given if the motif is not found within their promoter.

TRF2	Motifs Name	JASPAR id	Occurrence within 121 genes	Percentage Occurrence/121	Found TRF1 Promoter	Found TPP1 Promoter
1	SPIB	MA0081.1	100	82.64	1	1
1	NFATC2	MA0152.1	100	82.64	1	1
1	SP1	MA0079.3	101	83.47	1	1
1	ZNF354C	MA0130.1	94	77.69	0	1
1	KLF5	MA0599.1	95	78.51	1	1
1	ELF5	MA0136.2	82	67.77	1	0
1	VDR	MA0693.2	77	63.64	1	0
1	TFAP2A	MA0810.1	77	63.64	1	1
1	E2F6	MA0471.1	75	61.98	0	0
1	TFDP1	MA1122.1	73	60.33	1	0
1	TFAP2C	MA0524.2	71	58.68	1	1
1	TFAP2B	MA0811.1	67	55.37	1	1
1	Ddit3:Cebpa	MA0019.1	44	36.36	1	1

Table 9: Common motifs found in promoters of TRF2 and 119 downregulated genes.

The common motifs were then compared with the RNA-Seq data to see if the genes are being influenced in the presence or absence of LMP1 induction. Table 10 represents the results of the bioinformatic analysis combined with the RNA-seq results. The motifs found by using bioinformatics, are sites in which potential transcription factor proteins, listed in table 9, could potentially bind to our promoter region of interest. The genes were then verified in the RNA seq data to analyze and observe their behaviour when the cells are induced with LMP1. It was observed that SPIB, NFATC2, SP1 and Ddit3: CEBPA genes are upregulated, TFDP1 gene is downregulated in the presence of LMP1. KLF5 and E2F6 genes were found to be unaffected by the induction of LMP1 in the RNA-Seq data. ZNF354C, ELF5, VDR, TFAP2C, TFAP2B were not observed in the RNA-Seq data. These transcription factors have a probability of binding to our region of interest.

Motifs Gene Name	No LMP1 vs LMP1 Log2-ratio	RNA-Seq: RNA regulation in the presence of LMP1-Induction
SPIB	-0.9	Upregulated
NFATC2	-0.3	Upregulated
SP1	-0.3	Upregulated
ZNF354C	No Sample	No Data
KLF5	0.0	No Change
ELF5	No Sample	No Data
VDR	No Sample	No Data
TFAP2A	No Sample	No Data
E2F6	0.0	No Change
TFDP1	0.5	Downregulated
TFAP2C	No Sample	No Data
TFAP2B	No Sample	No Data
Ddit3:Cebpa	-0.7	Upregulated

Table 10: Motifs found using bioinformatics combined with the RNA-Seq data.
The RNA levels were analyzed in the presence and absence of LMP1 and annotated as upregulated, no change, no data or downregulated depending of its change in the cell when tetracycline was removed and LMP1 induced.

With this new data, DNA affinity chromatography was performed to determine what protein are bound to the region found between -130 bp and -100 bp of the TRF2 promoter and if any proteins mentioned above could be pull-down.

DNA Affinity Chromatography

The EMSA results determined protein complexes are binding to the 30 bp oligonucleotide probe of the TRF2 promoter region between -130 bp and -100 bp. The EMSA results showed different protein complexes binding to the 30 bp oligonucleotide probe in the nuclear extract with LMP1 induced and nuclear extract with no LMP1 induction. To determine what proteins were binding to the 30bp region, an oligonucleotide with a 5'-teg-biotin tag was ordered (Integrated DNA Technologies), in which the negative strand had the tag. (See protocol for DNA affinity chromatography mentioned in the materials and methods). By using streptavidin beads, the oligonucleotide was bound and with addition of 1.5 mg of nuclear extract the proteins were eluted through a series of washes. One-fifth of the sample was used for a gel analysis. A 10% acrylamide gel was used to separate the proteins and a silver stain was performed. As seen in Figure 35, lanes 1-3 are the inputs of the experiment and lanes 4-5 are the

proteins eluted. Lane 4 is the control with nuclear extract but no oligonucleotide. Lane 5 is the oligonucleotide with nuclear extract with no LMP1 induction; multiple bands are observed at different molecular weights. When nuclear extract from cells where LMP1 has been induced for 72hrs is used in the same conditions, the banding pattern changed and a new one appears, as observed in lane 6. Once the silver stain was obtained and a clearly different banding pattern was observed between the two experimental conditions, the remaining sample was sent to mass spectrometry.

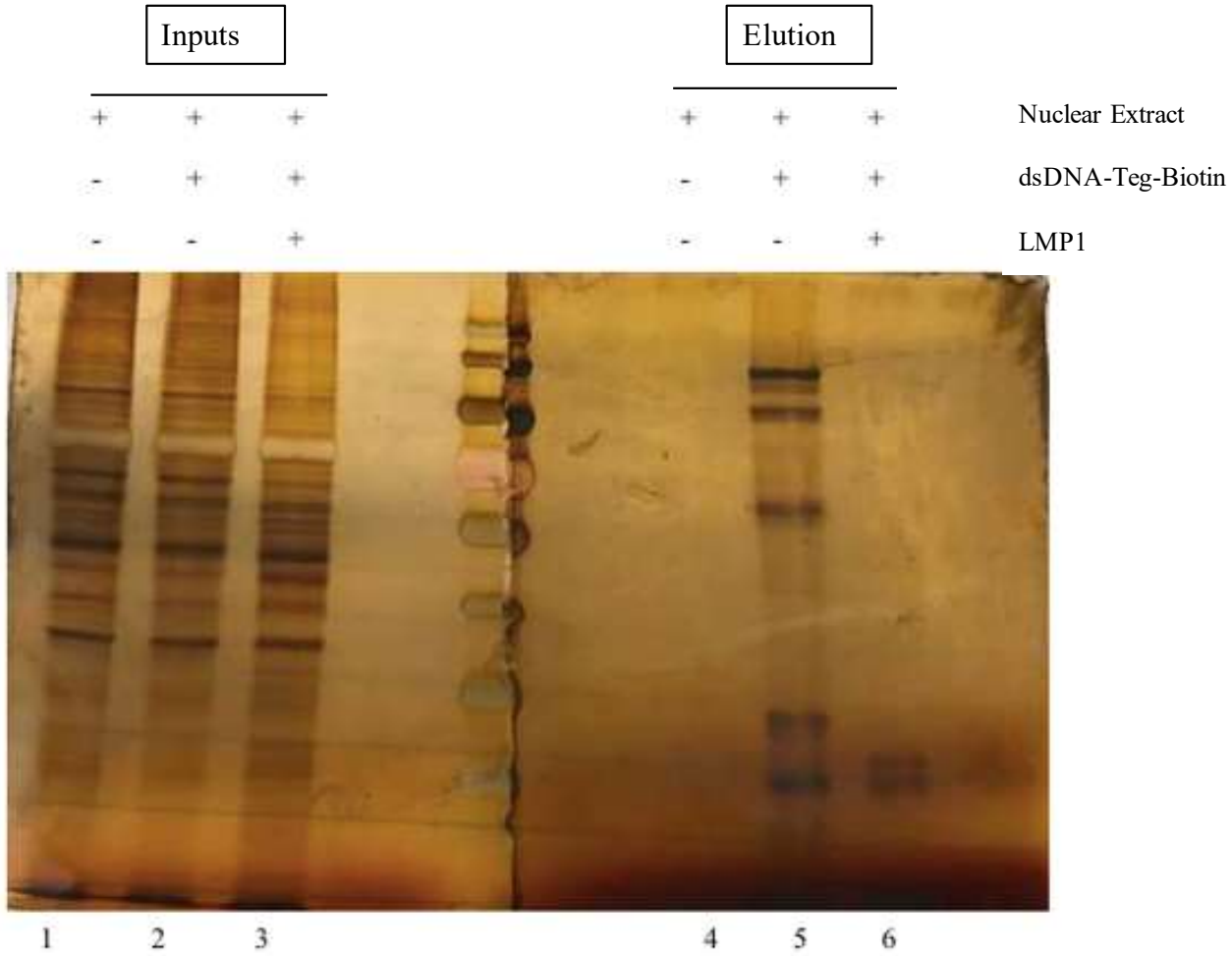


Figure 35: Silver stain from sample obtained after DNA affinity chromatography experiment.
A biotin labelled 5' end of the dsDNA oligonucleotide probe containing the region between -130bp and -100bp of the TRF2 promoter was used for this experiment. Lane 1 is input of negative control with no oligonucleotide, Lane 2 is input of nuclear extract non-induced with LMP1, Lane 3 is input of nuclear extract of LMP1-induction, Lane 4 is the eluted negative

control, lane 5 is the elution when a nuclear extract with no LMP1 and Lane 6 is with LMP1. Different band patterns appear in Lane 6 when an LMP1-induced nuclear extract was used.

Mass spectrometry was performed by the laboratory of Francois-Michel Boisvert to determine what was bound to the 30 bp region of interest. 1169 proteins were identified. See annex Table 15-16 for the complete list of proteins identified. A negative control (with no oligonucleotide dsDNA) was used to eliminate any non-specific proteins binding to the streptavidin beads. All samples that had proteins lesser or equal to the negative control found during mass spectrometry analysis were dismissed. Afterwards, the samples with LMP1-induction versus the sample with No LMP1 were compared to each other to find which had more proteins in one compared to the other. What we want is to find specific proteins in each sample. If the total protein MS/MS count was lower than 2 peptides, the proteins were not considered further. If the proteins between the LMP1 non-induced and induced samples were equal to each other they were grouped together for potential further analysis. The proteins were then separated by the protein MS/MS counts and assigned to two separate groups. The groups were determined as follows: proteins where LMP1>No LMP1 were assigned to Group "LMP1"; and proteins where No LMP1>LMP1 were assigned to Group "No LMP1". Once the proteins were separated into the "No LMP1" and "LMP1" groups, a ratio between the Total MS count over the sample was made between the two samples. Then a cut-off of 2-fold increase was made. This is to determine what proteins within the samples were found to be 2-times greater than the sample being compared to. The experiment and analysis was repeated a second time using the same parameters. Once obtained, the samples N=1 and N=2 for both "No LMP1" and "LMP1" were verified for a common set of proteins. For the No LMP1 experiment, it was found that N=1 had 434 proteins, N=2 had 371 proteins, and there were 293 proteins in common between the two assays (reproducible by 73%). For the LMP1 experiment, it was found that N=1 had 227 proteins, N=2 had 198 proteins, and there were 144 proteins in common between the two assays (reproducible by 66%). The set of proteins found in each experiment were taken as the group of specific proteins for further analysis (Figure 36).

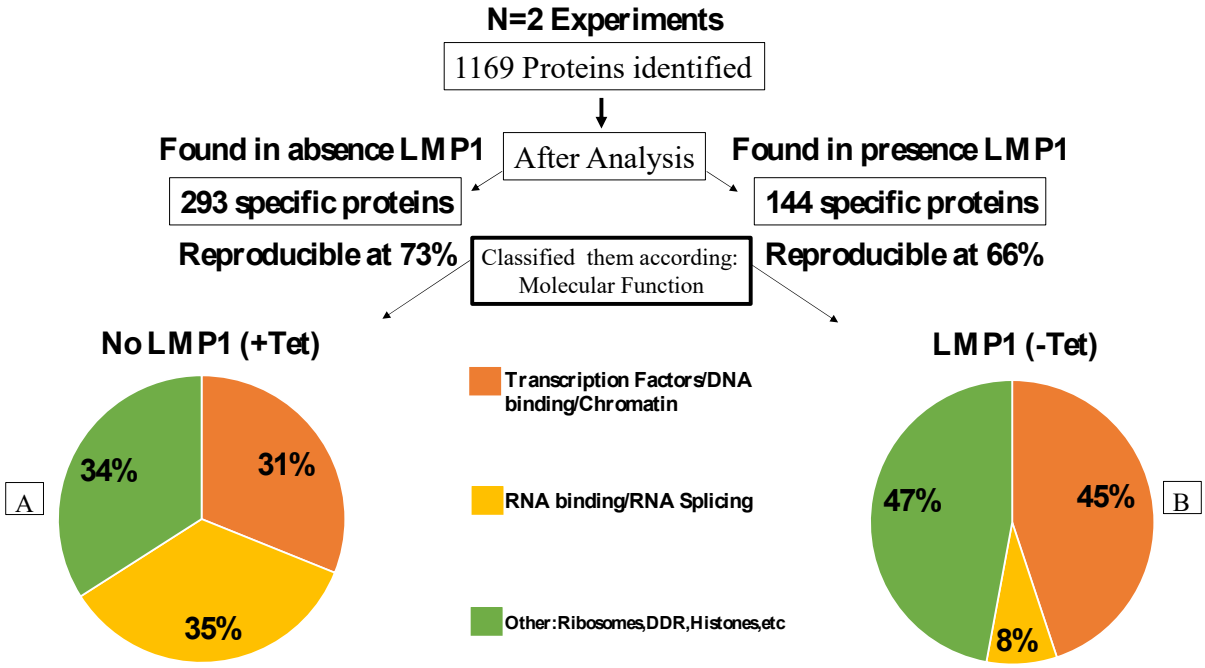


Figure 36: Schematic chart analysis of the Mass Spectrometry results of No LMP1 versus LMP1 samples from DNA Affinity Chromatography experiments.
Pie charts were designed to show to percentage of proteins that went into each category. A) Sample from the no LMP1(+Tet) in which 31% proteins fall in the category of transcription factors, core promoter binding, chromatin binding, sequence-specific DNA binding, protein in the basal transcription machinery, 35% are RNA binding or RNA splicing and 34% are other proteins such as ribosomes, DNA-Damage response proteins, histones and etc. B) are samples from the LMP1-induced (-Tet) in which 45% proteins fall in the category of transcription factors, core promoter binding, chromatin binding, sequence-specific DNA binding, protein in the basal transcription machinery, 8% are RNA binding or RNA splicing and 47% are other proteins such as ribosomes, DNA-Damage response proteins, histones and etc.

For the full list of proteins found in the MS analysis, please see Annex 14-15. Proteins were classified according to their molecular function as per Go Annotation: Gene ontology Consortium (http://www.geneontology.org/). The proteins were then chosen for analysis based on their unique peptides count, peak intensities, and function. In parallel, they were verified against the RNA-Seq experiment to compare if they are upregulated or downregulated in the presence of LMP1. The proteins shown below, table 12-13, are the top ten proteins that were found within the categories of core promoter binding or sequence-specific DNA binding for the No LMP1 and LMP1 group. The region being studied is a small region of the TRF2 promoter; therefore, we wanted to focus on proteins possibly binding to DNA or part of the basal transcription machinery.

Table MS No Change Between No LMP1 and LMP1

Protein Name	Unique Peptides	Sequence Coverage	Peak Intensities	RNA Seq	GO Annotation Molecular Function
SPIB	4	11	1.15E+08	Upregulated	Sequence-specific DNA binding
NFATC2	5	3.6	3.30E+06	No Change	Sequence-specific DNA binding
Pax5	9	44.2	7.09E+09	No Change	Sequence-specific DNA binding
SPI1	2	28.1	3.66E+08	Downregulated	Sequence-specific DNA binding

Table 11: Proteins found in both No LMP1 and LMP1 sample with no apparent change in the mass spectrometry analysis and are sequence-specific DNA binding proteins.

The bioinformatics found possible proteins that could be binding to the region of interest that is between -150bp and -100bp. With the mass spectrometry analysis, we were able to pull-down proteins previously mentioned by the bioinformatics identified to bind to conserved regions. SPIB and NFATC2 were proteins pull-down with the 30bp region and identified by mass spectrometry (Table 11). These proteins were found in both samples in the presence or absence of LMP1. I also performed promoter sequence analysis to find possible transcription factors binding to our region. Pax5 and SPI1 were two different proteins identified from the 14 others identified in table 8. Pax5 and SPI1 were found by mass spectrometry analysis (Table 11). These proteins are candidates for binding to our region and possibly being regulated by unknown mechanism during LMP1-induction.

Table MS No LMP1

Protein Name	Unique Peptides	Sequence Coverage	Peak Intensities	RNA Seq	GO Annotation Molecular Function
DNMT1	51	36.5	2.05E+09	No Change	Sequence-specific DNA binding
HNRNPU	45	45.9	3.02E+09	No Change	Core promoter binding
IFI16	22	50.2	8.24E+09	Upregulated	Core promoter binding
MED1	19	15.1	3.40E+08	No Change	Core promoter binding
BRD4	17	14	7.88E+08	Upregulated	Core promoter binding
RUVBL2	17	29.4	3.28E+08	Downregulated	Core promoter binding
SLTM	16	17	8.42E+08	Upregulated	Sequence-specific DNA binding
BRD2	14	29.7	4.79E+08	Upregulated	Core promoter binding
THRAP3	12	16.4	8.03E+08	No Change	Core promoter binding
BRD3	8	27	3.95E+08	No Change	Core promoter binding

Table 12: Top 10 proteins found in mass spectrometry for sample with No LMP1 induction that are either core promoter binding proteins or sequence-specific DNA binding proteins.

From the DNA affinity chromatography and mass spectrometry, multiple proteins with different roles were identified. The list in table 12, are the top proteins that were assigned, by Go Annotation, a category such as core promoter binding or sequence-specific proteins. Many proteins play a role in chromatin remodelling such as DNMT1, BRD2, BRD3, BRD4 and RUVBL2. BRD4 plays a key role in regulating the transcription of signal-inducible genes by associating with the P-TEFb complex and recruiting it to promoters (Ai et al. 2011) IF116 is a transcription activator binding to double-stranded DNA (Li et al. 2012) and a potential candidate for further analysis. The MED1 was also found through mass spectrometry, a co-activator protein important in transcription regulation. A lot of proteins found in the No LMP1 data are

interesting proteins as they a coincides with transcription DNA-binding or chromatin remodelling.

When it came to protein data found in the LMP1 samples analyzed by mass spectrometry in the region between -130 bp and -100bp of the TRF2 promoter, the protein composition changed (Table 13). Interestingly, two repressors were found: CTCF and YY1;YY2 (Austen, Luscher & Luscher-Firzlaff, 1997) MYBBP1A also plays a role in repression when interacting with sequence specific DNA-binding proteins (Hara et al., 2009).

Table MS LMP1 (-Tet)

Protein Name	Unique Peptides	Sequence Coverage	Peak Intensities	RNA Seq	GO Annotation Molecular Function
UBTF	45	43.3	1.09E+10	No Change	Core promoter binding
WDR76	19	27.5	7.40E+08	No Change	Core promoter binding
NOC2L	19	21.9	1.42E+09	No Change	Sequence-specific DNA binding
CTCF	17	18.8	1.77E+09	Downregulated	Sequence-specific DNA binding
MYBBP1A	13	7.8	1.86E+08	No Change	Core promoter binding
ZNF106	12	17.9	2.65E+08	No Change	Sequence-specific DNA binding
ZNF629	10	10.4	3.54E+08	No Change	Sequence-specific DNA binding
YY1; YY2	10	13.5	1.48E+09	No Change	Core promoter binding
ABT1	6	20.2	1.97E+08	No Change	Sequence-specific DNA binding
HOXB4	6	17.9	2.37E+08	No Change	Sequence-specific DNA binding

Table 13: Top 10 proteins found in mass spectrometry in sample induced with LMP1 that are either core promoter binding proteins or sequence-specific DNA binding proteins.

Discussion and Conclusions

EBV-positive Hodgkin lymphoma has been shown to have telomeric dysfunction as cells transition from mononucleated cells into multinucleated cells. It was also published by Lajoie et al., 2015, that the oncogenic protein LMP1 produces multinucleation and downregulates TRF1, TRF2 and POT1 through pathways that remain to be determined. Further, LMP1 had no effect on TRF2 regulation when a Myc-driven TRF2 was reintegrated into a BJAB tTA LMP1 cell line expressing LMP1. These experimental analyses therefore shed light into how a virally induced cancer originates via the regulation of telomere-related proteins, in particular TRF2. It was then hypothesized that LMP1 might play a role in the downregulation of TRF2 by affecting its promoter function. Understanding how TRF2 is expression is regulated might allow us to determine a new pathway of regulation by LMP1 in Hodgkin Lymphoma.

LMP1-induction affects TRF2 at its promoter

LMP1 is a trans-membrane protein that mimics a cellular necrosis receptor (TNFR) family member. Its two spanning domains, referred to as CTAR-1 and CTAR-2, have been shown to be critical for signalling through TNFR-associated factors (TRAFS) (Devergne et al., 1996), activation of transcription factors such as NF-κB (Brodeur, Cheng, Baltimore & Thorley-Lawson, 1997), AP-2 (by p38 mitogen-activated protein kinase) (Eliopoulos et al., 1999) and AP-1 via the JNK pathway (Eliopoulos, Blake et al., 1999). As a key factor in B-cell transformation, the pathways by which it activates and inhibits transcription factors are not fully understood. It was shown by Lajoie et al., 2015, that if a myc-driven TRF2 is integrated in an LMP1-expressing cell, this expression did not affect the promoter function. We therefore hypothesize that LMP1-induction might be interfering with TRF2 expression via gene promoter inhibition. In previous studies, it was shown by Everly, Mainou and Raab-Traub, 2009, that LMP1 decreases p27 RNA levels and the activity of the p27 promoter. LMP1 was found to regulate the promoter element through the E2F family transcription factors (E2F4 and p130). Therefore, to determine if LMP1 affects TRF2 via the promoter region, constructs were made in which a GFP reporter gene was placed under the control of various portions of the TRF2 promoter. The precise constitution of the TRF2 promoter is unknown. Furthermore, there is a lack of knowledge regarding its regulation in terms of

silencers or other regulatory sites. For the dissection of the TRF2 promoter, we started with 2566bp and 1307 bp of upstream sequences from the translation start site (TSS) of the ATG (+1), including the known 5'UTR of 17bp. This length was chosen in order to try to include any regulatory sites, as they can be found thousands of base pairs away from the gene. It was important to include the 5' UTR, as it can also be regulated by different mechanisms. By Western blot, we showed that the GFP protein is downregulated with both the 2566bp and 1307 bp TRF2 promoter fragments (Figure 17). To our knowledge, these findings provide the first demonstration that LMP1-induction can influence the regulation of a telomeric protein, TRF2, via its promoter sequences.

Flow cytometry is another technique that we attempted to use to determine the GFP downregulation as observed by Western blot. This approach was thought to be simple and rapid, as no antibody was needed, and the protein can be detected directly inside cells. To start the analysis, a negative control was created with no promoter (PL) upstream the GFP gene and analyzed via flow cytometry to determine where our peaks fall. Figure 18 represents the data obtained for the cells expressing or not LMP1. This represents the background noise for when cells were transfected with the PL-GFP. We also included a positive control with a CMV promoter as shown in Figure 19. As predicted, we found the plasmid construct with the TRF2 promoter in between the two controls, as seen in Figure 20 and 21. However, the analysis could not be continued with flow cytometry as the plasmids, p2566-GFP and p1307-GFP did not show any detectable and reliable shift when LMP1 expressing cells were compare to non-expressing cells. One explanation for this may be that TRF2 has a weak promoter, and therefore weak GFP protein expression. When LMP1 is expressed, the TRF2 promoter is downregulated by approximately 40% (Lajoie et al 2015), therefore, if we have a weak promoter and a small amount of GFP protein produced, the flow cytometry might not be sensitive enough to detect the change.

Another way to assess eGFP gene expression, could be real-time PCR (RT-qPCR). For these analyses, different primers were used for quantification as shown in the materials and methods section. Some difficulties arose when analyzing the data in Figure 22. The negative control, positive control and the experimental samples all had similar high CT values of over 30.

A ct value above 30 indicates minimal amounts of the target nucleic acids being amplified. However, the housekeeping genes were located in the range of 23-24 for CT values, which is within the expected range for successful RT-qPCR target amplification, indicating that the problem was specifically with amplification of the GFP mRNA target. The melting curves showed one peak for each primer, thus confirming that it was targeting one amplicon and not many. The amplicons were also verified on agarose gel. There was amplification of the no cDNA template control in the mixture, indicating possible contamination by GFP. To eliminate the possible contamination and to improve reproducibility, RT mix (Quantitect RT Kit 205311) Biorad qPCR buffer (iScript^{TM}RT supermix, Product 1708840) was used. In order to optimize the RT-qPCR, different concentrations of the cDNA for qPCR were quantified. By performing a serial dilution of a sample containing cDNA for qPCR, it was found that for the primers used in this study, a good starting amount was 2ug of cDNA to obtain sufficient amplification during qPCR. A standard primer dilution curve was made to test the efficiency of the primers at different concentrations. These modifications and optimizations led to better results for the housekeeping genes, TRF2 (Figure 28) and LMP1 (Figure 27), however the eGFP expression levels were still not comparable, as only low expression of the target gene was detected (Figure 25-26). Also, when qPCR samples were loaded on a gel, the no template sample was still amplifying the GFP gene. To resolve this problem, it was decided to utilize the relatively new technique ddPCR for eGFP analysis, as it has shown to be beneficial when trying to quantify low abundance targets. In addition, to reduce the possibility that contamination could be coming from our lab, after cells were harvested, they were mixed with Trizol and given to the Université de Sherbrooke RNomics platform for RNA extraction, RT and ddPCR analysis. Furthermore, new primers targeting the end of the gene were designed for the eGFP quantification. Since the eGFP gene does not have an exon, the end of the gene was targeted for better amplification and also to minimize any non-specific annealing with other genes. The primers GFP F16R16 and GFP F18R18 were used in the ddPCR analysis. As can be seen in both Figures 25-26 of the eGFP analysis by ddPCR, the plasmid with no promoter upstream the eGFP gene is being amplified within the samples. Samples that could not be detected in the Western blot analysis seemed to be present in the ddPCR, such as the promoterless plasmid, and the plasmids containing 50 bp and 100 bp promoter regions. Even more concerning, the plasmids that expressed detectable eGFP in our Western blot, seemed to be comparable in terms of RNA expression to the promoterless

plasmid. Thus far, we have found it difficult to explain why eGFP is detectable in ddPCR, other than the possibility there is GFP contamination in the buffers or RT mix being used in RT-qPCR experiment. Upon discussion with other laboratories using GFP as a reporter gene, unexplained GFP signal was found to be a common and recurring problem. For future experiments, the construct could be switched to a luciferase gene in which both RT-qPCR and fluorescent analysis could be performed. This could confirm RNA expression levels in the experiment in which a downregulation can be seen during LMP1 induction. However, for the purposes of this project, we decided to continue analyzing GFP-expression from the various promoter constructs via Western blot analysis as the plasmid constructs were already done and a downregulation could be documented with this approach.

Identifying a smaller region of TRF2 affected by LMP1-induction

By Western blot analysis (Figure 17) it was determined that LMP1 induction downregulated expression of GFP via the TRF2 promoter. To determine a region of regulation, the TRF2 promoter was progressively shortened until a region would no longer be affected by LMP1-induction. We started with regions containing 2566bp and 1307 bp upstream of the ATG and then dissected with shorter promoters. The first constructs used for analysis were 821bp, 737bp, 680bp, 614bp and 484 bp all of which showed a downregulation in the presence of LMP1-induction. The downregulation for promoters 680bp, 614bp was not as pronounced as the other promoters. Two things will influence the experiment: level of transfection of the plasmid into the cell and the induction level of LMP1 upon removal of tetracycline. Therefore, depending on the experiment and the cells' ability to express LMP1, it will influence and vary the degree of downregulation on the promoters. The experiment was repeated three times, confirming that LMP1-induction downregulated the TRF2 promoter of 821bp, 737bp, 680bp, 614bp and 484 bp (Figure 29). We then constructed plasmids with smaller promoter regions comprising 400bp, 300bp, 250bp and 200bp. Again, the downregulation in the presence of LMP1- induction persisted in all these constructs (Figure 30). Therefore, the site of regulation on the TRF2 promoter should be further downstream. For further analysis, constructs of 150bp, 100bp and 50 bp were made. As seen in Figure 31, in the construct of 150bp, a good GFP expression can be observed by Western blot but when a length of 100 bp promoter was used, a low level of GFP

expression was seen. If a 50bp promoter length construct was used, minimal to no expression could be detected. This analysis indicated that a promoter of minimally 100 bp is needed for proper expression. We hypothesize that the regulation at the TRF2 promoter is occurring between 150bp and 100bp. To find a smaller region for analysis, the TRF2 promoter was dissected between the 150bp and 100bp and constructs of promoter length of 140bp, 125bp and 115bp were made for further analysis. In fact, 140bp as well as 125bp were still enough for the downregulation by LMP1-induction (Figure 32). When a promoter length of 115bp was used, only a low amount of eGFP can be detected by Western blot analysis and a downregulation could still be observed. It was reasoned that a minimum of 125 bp of the TRF2 promoter was needed for gene transcription. At 125 bp, downregulation affected by LMP1-induction could be easily observed. In parallel, the endogenous TRF2 was also probed with anti-TRF2 to show that it was also downregulated by the presence of LMP1-induction in the cell. Quantification was also done on three independent experiments to visualize the GFP expression of promoters 150bp, 140bp, 125bp, 115bp, 100bp and the promoterless via a bar graph (Figure 33). Only one quantification was done for LMP1 and TRF2 and this was taken from the Western blot in Figure 32. Figure 33 visually represents the downregulation occurring between promoter 150bp-100bp when LMP1 is induced. Even if the expression of GFP is minimal on a Western blot, a downregulation by LMP1 is still observed down to 100bp. To move forward with the TRF2 promoter analysis, it was decided to take the region between -130bp and -100bp for protein analysis.

RNA Sequencing Analysis

RNA Seq was performed on the cells cultivated for 72hrs after addition or removal of tetracycline for the induction of LMP1. This technology was used to provide information on all of the transcripts that were upregulated or downregulated. For our specific case, we wanted to assess what other genes besides TRF2 were downregulated upon LMP1 induction. Data analysis of the RNA-Seq results confirmed that 119 genes had equal to or greater downregulation than TRF2. With these new genes found to be regulated in the presence of LMP1, a bioinformatic analysis was performed to detect possible common motifs in the region of -130 bp and -100 bp of the TRF2 promoter vs the region of 500bp of each of the promoters of the co-regulated genes. With this approach, 14 motifs were found, as described in table 9. The analysis was also

compared with TPP1, since this gene was not affected by LMP1-induction. It was found that ELF5, VDR E2F6 and TFDP1 motifs were found in both the TRF1 and TRF2 promoters but not the TPP1 promoter. Therefore, the RNA Seq provided a starting point as to we might expect as factors to be found via DNA Affinity Chromatography and Mass Spectrometry analysis.

RNA seq did confirm some upregulated genes that play role in the LMP1 signalling pathway, such as TRAF1, STAT3, STAT2, STAT5A, STAT1, ATF2 and JUNB. When these genes are activated, they induce numerous downstream effectors and impact a variety of cellular processes such as proliferation, cytokine induction, apoptosis and cell survival (Morris, Dawson & Young, 2009). These data hence confirm that our RNA Seq is comparable to other published articles showing genes affected by LMP1-induction.

Protein Analysis on promoter region between -130bp and -100bp

To determine protein-DNA interactions, an EMSA was performed. A 5' γ-ATP^{32} radiolabelled 30bp dsDNA probe corresponding to the region between -130bp and -100bp of the TRF2 promoter and nuclear extract from LMP1-expressing or non-expressing cells were used for the assay. After including specific and non-specific competitor DNA in the assay, it could be determined that we had a number of specific complexes binding to our 30bp probe. When the experiment was repeated using a nuclear extract in which LMP1 was induced for 72hrs, there was a significant change in observed banding patterns. This confirmed that LMP1-induction creates a cascade of events leading to a change in protein composition within the nucleus. Altogether, this EMSA assay confirmed two things: firstly, we have protein-DNA interaction with our probe and secondly, we have a different protein composition binding to our probe depending on the presence or absence of LMP1 in the cell.

To determine what the different proteins were on the small region of the TRF2 promoter, a pull-down was performed (also known as DNA affinity chromatography). The BJABJ tTa LMP1 cell line was used to obtain nuclear extracts with and without LMP1. By using a biotin-teg tag on a dsDNA 30bp region (-130bp and -100 bp), and nuclear extract, proteins were pull-down and analyzed by PAGE. The silver stained gel revealed multiple bands with both nuclear

extracts. However, the patterns of them were very different with many bands in the absence of LMP and two major bands that appear around 26kDa with LMP1 (Figure 35). These results are congruent with the EMSA results, confirming there are different protein interactions within the 30bp region in the presence or absence of LMP1.

Finally, the proteins binding to the region between -130 bp and -100bp were identified by mass spectrometry. After systematic analyses, the data were placed in two groups: No LMP1 (cells with tetracycline) and LMP1 (cell induced by removal of tetracycline). By doing so, 293 specific proteins were found in the No LMP1 group and 144 specific proteins were found for the LMP1 group. The reproducibility for No LMP1 was 73% and 66% for LMP1. Table 10 and Table 11 represent the top 10 proteins found in both groups. For the No LMP1 group, interestingly, the top protein found was DNMT1 and that this protein was not found within the samples with LMP1. DNA methyltransferase (DNMT) is a common epigenetic factor that methylates CpG islands and contributes to regulation of gene expression (Robertson, 2001). Upon LMP1-induction, DNMT1 and DNMT3b were found to be downregulated in GC B cells (Leonard et al. 2011) This is significant as our RNA seq data also confirmed that both DNMT1 and DNMT3b are downregulated within the data set. Dnmt1 also interacts with a large number of repressor proteins, such as histone deacetylases to inhibit transcription in a methylation-independent manner. Based on this information, DNMT1 might not be the protein of interest in our case as it is a transcriptional repressor while our results require a transcriptional activator.

One of the top 3 proteins found in mass spectrometry in the group No LMP1 was a protein called IFI16. The IFI16 was found to be upregulated in the RNA seq data. It is a member of the interferon-inducible HIN-200 family of nuclear proteins that is implicated in activation in transcriptional regulation by modulating protein-protein interactions. IFI16 is involved in p53-mediated transcriptional activation and apoptosis by upregulating p53 and downregulating MDM2 (Fujiuchi et al., 2004). These data coincide with other data that shows LMP1 suppressing MDM2 and thus promoting p53 accumulation (Li et al., 2012) in which in this case, could be because of IF116 upregulation. This is interesting because our region of interest is a double-stranded DNA promoter region and we are trying to identify regulatory proteins binding to this region. Since it is upregulated in the presence of LMP1-induction as shown in the RNA-seq data,

it could be hypothesized that it could be binding to our promoter region but its protein partners or mediators are downregulated, therefore, decreasing interaction and possible recruitment of the transcription machinery. This would in fact decrease the expression levels of TRF2 during transcription.

Combined with the bioinformatics analysis, 14 different motifs were found within the 30bp region of the TRF2 promoter and these motifs were also found in 82% of the 119 genes co-downregulated. Mass spectrometry of the No Change group, found SPIB, and NFATC2 that were pulled-down and the binding motifs of these proteins are included in the 14 motifs above. However, these proteins were grouped separately as they were found equally in both No LMP1 and LMP1 samples. SP1I and Pax5 were also added as they were found using ALGGEN.com program (identification of putative transcription binding sites in DNA sequences) (Table 9). To fit with our hypotheses, MED1 and THRAP3 (Table 10) could be potential candidates as both are core promoter binding proteins and part of basal transcription machinery. Not shown in the top proteins but can be found in the annex, are other MED proteins, ERCC2/ERCC3 and also transcription factors such as GTF2 (H4, F1, H2C, H2, E2), and MNAT1. These proteins are all part of basal transcription machinery. For the LMP1 group, the entire protein composition changes, as confirmed by EMSA and silver stain, instead of having the basal transcription factors, DNMT1 or IFI16. The shift is towards activator or repressor proteins such as YY1; YY2 and CTCF, which are known gene regulators (Austen et al., 1997) (Table 11).

RNA Seq analyses combined with mass spectrometry will need to undergo further refinement. However, a preliminary hypothesis emerges from this information. It is possible that the proteins found by bioinformatics and mass spectrometry (Table 9) are proteins that are bound to the promoter in the presence or absence of LMP1. These could be the key proteins that interact with the DNA and then recruit the transcription machinery to the promoter. We further speculate that the proteins found in the No LMP1 group could be found in a complex(es) and interact with the DNA bound proteins on the promoter to initiate transcription. Therefore, it is also possible that these proteins such as SPIB and NFATC2 are stabilized by its interacting partners to the DNA and recruit the transcription machinery. Then upon LMP1-induction, their interacting partners are downregulated, thus decreasing transcription of TRF2. If SPIB, Pax5 or NFATC2 need an

interacting partner to stay bound at the promoter, but during LMP1-induction this partner is downregulated, this would cause the proteins to dissociate from the DNA. Then during LMP1-induction, it would give the chance of competing proteins to associate with the promoter. In this case, there will be site competition and giving the opportunity of other regulatory repressor proteins such as CTCF or YY1; YY2 proteins to bind to the promoter (Table 11). These proteins were found in the mass spectrometry in our LMP1 group, giving new insight that a repressor might be considered in the TRF2 promoter regulation during LMP1-induction. These hypotheses need to be tested in future experiments.

The results of this project clearly demonstrate that the LMP1 mediated signalling pathway interferes with TRF2 transcription at its promoter. To date, only one EMSA was performed using nuclear extracts using the same extracts and then used for mass spectrometry. To confirm the results, these experiments will be repeated with new nuclear extracts for both the EMSA and mass spectrometry. Additional EMSAs will be performed using BJAB tTA cells with the addition and removal of tetracycline to verify that the tetracycline is not the cause of the protein shifts. Even though the proteins previously mentioned are proteins with the function of DNA binding, chromatin DNA binding and RNA splicing proteins cannot be excluded as they all could play a dynamic role in transcription. As many genes were found to be upregulated or downregulated as well as the different proteins found during the mass spectrometry analysis, further work will be needed to elucidate the role of LMP1-induction at the TRF2 promoter. By understanding how LMP1-induction affects the regulation of TRF2 at its promoter, we could determine a pathway back the trans-membrane LMP1 to fully understand its role in the cascade of events leading to the downregulation of TRF2.

References

Austen, M., Luscher, B., & Luscher-Firzlaff, J.M. (1997). Characterization of the transcriptional regulator YY1. *Journal of Biological Chemistry, 272*(3)*,* 1709-1717. doi:10.1074/jbc.272.3.1709

Babcock, G., Hochberg, D., & Thorley-Lawson, D.A. (2000). The expression pattern of Epstein-Barr virus latent genes in vivo is dependent upon the differentiation stage of the infected B cell. *Immunity, 13*(4), 497-506. http://doi.org/10.1016/S1074-7613(00)00049-2

Badaracco G., Bianchi, M., Valsasnini, P., Magn, G., & Plevani, P. (1985). Initiation,elongation and pausing of in vitro DNA synthesis catalyzed by immunopurified yeast DNA primase: DNA polymerase complex. *EMBO Journal, 4*(5), 1313-1317.

Bailey, S.M., & Murname, J.P. (2006). Telomeres, chromosome instability and cancer.*Nuclei Acids Research, 34*(8), 2408-2417. http://doi.org/10.1093/nar/gkl303

Balakrishman, L., & Bambara, R.A. (2013). Okazaki fragment metabolism. *Cold Spring Harbor Perspectives in Biology*, *5*(2). doi:10.1101/cshperspect.a010173

Benarroch-Popivker, D., Pisano, S., Mendez-Bermudez, A., Lototska, L., Kaur, P., Bauwens, S., Giraud-Panis, M.J. (2016). TRF2-mediated control of telomere DNA topology as a mechanism for chromosome-end protection. *Molecular Cell, 61*(2), 274-286.

Boyd, J.M., Gallo, G.J., Elangovan, B., Houghton, A.B., Malstrom, S., Avery, B.J., … Lutz, R.J.,et al. (1995). Bik, a novel death-inducing protein shares a distinct sequence motif with Bcl-2 family proteins and interacts with viral and cellular survival-promoting proteins. *Oncogene*, *11*(9), 1921-1928.

Brauninger, A., Schmitz, R., Bechtel, D., Renne, C., Hausmann, M.L., & Kuppers R. (2005). Molecular biology of Hodgkin's and Reed/Sternberg cells in Hodgkin's lymphoma. *International Journal of Cancer, 118*(8), 1853-1861.

Brodeur, S. R., Cheng, G., Baltimore, D., & Thorley-Lawson, D.A. (1997). Localization of the major NF-κB-activating site and the sole TRAF3 binding site of LMP-1 defines two distinct signaling motifs. *Journal of Biological Chemistry*, *272,* 19777-19784.

Broustas, C.G., & Lieberman, H.B. (2014). DNA damage response genes and the developmentof cancer metastasis. *PMC 181*(2), 111-130. doi:10.1667/RR13515.1

Chan, S.W., & Blackburn, E.H. (2002). New ways not to make ends meet: telomerase, DNA damage proteins and heterochromatin. *Oncogene, 21*(4)*,* 553-563.

Ciccia, A., & Elledge, S.J. (2010). The DNA damage response: Making it safe to play withknives. *Molecular Cell, 40*(2), 179-204. doi:10.1016/j.molcel.2010.09.019

Cohen, S.B., Graham, M.E., Lovrecz, G.O., Bache, N., Robinson, P.J., & Reddel, R.R. (2007). Protein composition of catalytically active human telomerase from immortal cells. *Science, 315*(5820), 1850-1853.

Cong, Y.S., Wright, W.E., & Shay, J.W. (2002). Human telomerase and its regulation. *Microbiology and Molecular Biology Reviews, 66*(3), 407-425. doi:10.1128/MMBR.66.3.407-425.2002

d'Adda di Fagagna, F., Reaper, P. M., Clay-Farrace, L., Fiegler, H., Carr, P., Von Zglinicki, T., Jackson, S.P. (2003). A DNA damage checkpoint response in telomere-initiated senescence. *Nature, 426*, 194-198.

Davoli, T., & de Lange, T. (2012). Telomere-driven tetraploidization occurs in human cells undergoing crisis and promotes transformation of mouse cells. *Cancer Cell, 21*(6), 765-776.

Dawson, C., Tramountanis, G., Eliopoulos, A.G., & Young, L.S. (2003). Epstein-Barr virus latent membrane protein 1 (LMP1) activates the phosphatidylinositol 3-kinase/akt pathway to promote cell survival and induce actin filament remodeling. *Journal of Biological Chemistry, 278*(6), 3694-3704.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards humatelomeres. *Genes and Development, 19*(18), 2100-2110. doi:10.1101/gad.1346005

Devergne, O., Hatzivassiliou, E., Izumi, K.M., Kaye, K.M., Kleijnen, M.F., Keff, E., & Mosialos, G. (1996). Association of TRAF1, TRAF2, and TRAF3 with an Epstein-Barr virus LMP1 domain important for B-lymphocyte transformation: role in NF-κB activation. *Molecular and Cellular Biology, 16*(12), 7098-7108.

Dimitrova, N., & de Lange, T. (2009). Cell cycle-dependent role of MRN at dysfunctionaltelomeres: ATM signaling-dependent induction of nonhomologous end joining (NHEJ) in G1 and resection-mediated inhibition of NHEJ in G2. *Molecular and Cellular Biology, 29*(20), 5552-5563. doi:10.1128/MCB.00476-09.

Eliopoulos, A.G., Blake, S.M., Floettmann, J.E., Rowe, M., & Young, L.S. (1999). Epstein-Barr virus-encoded latent membrane protein 1 activates the JNK pathway through its extreme C terminus via a mechanism involving TRADD and TRAF2. *Journal of Virology, 73*(2), 1023-1035.

Eliopoulos, A.G., Gallagher, N.J., Blake, S.M., Dawson, C.W., & Young, L.S. (1999). Activation of the p38 mitogen-activated protein kinase pathway by Epstein-Barr virus-encoded latent membrane protein 1 coregulates interleukin-6 and interleukin-8 production. *Journal of Biological Chemistry, 274*(23), 16085-16096.

Everly, D.N., Mainou, B.A., & Raab-Traub, N. (2009). Transcriptional downregulation of p27KIP1 through regulation of E2F function during LMP1-mediated transformation. *Journal of Virology, 83*(24), 1271-12679. doi:10.1128/JVI.01422-09

Floettmann, J.E., Ward, K., Rickinson, A.B. & Rowe, M. (1996). Cytostatic effect of Epstein Barr virus latent membrane protein-1 analyzed using tetracycline-regulated expression in B cell lines. *Virology, 223*(1), 29-40.

Freshney, R.I. (1994). *Culture of animal cells: A manual of basic technique* (3rd ed.). NeYork: Wiley-Liss Inc.

Frick, D.N., & Richardson, C.C. (2001). DNA primases. *Annual Review of Biochemistry, 70*,39-80.

Fuks, F., Bergers, W.A., Brehm, A., Hughes-Davies, L., & Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genetics, 24*(1), 88-91.

Fujiuchi, N., Aglipay, J.A., Ohtsuka, T., Maehara, N., Sahin, F., Su, G.H., … Ouchi, T. (2004). Requirement of IFI16 for the maximal activation of p53 induced by ionizing radiation. *Journal of Biological Chemistry 279*(19), 20339-20344.

Gahn, T., & Sugden, B. (1995). An EBNA1 dependent enhancer acts from a distance of 10 kilobase pairs to increase expression of the Epstein-Barr virus LMP gene. *Journal of Virology, 69*(4), 2633-2636.

Ganti, S.N., Albershardt, T.C., Iritani B.M., & Ruddell, A. (2015). Regulatory B cells preferentially accumulate in tumor-draining lymph nodes and promoter tumor growth. *Scientific Reports, 5,* 12255.

Gathings, W.E., Lawton, A.R., & Cooper, M.D. (1977). Immunofluorescent studies of the development of pre-B cells, B lymphocytes and immunoglobulin isotype diversity in humans. *European Journal of Immunology, 7*(11), 804-810. doi:10.1002/eji.1830071112

Gobbi, G. P., Ferreri, J.M.A., Ponzoni, M., & Levis, A. (2013). Hodgkin lymphoma. *Critical Reviews in Oncology/Hematology*, *85*(2), 216-237. doi:10.1016/j.critrevonc.2012.07.002

Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Cell Biology, 89*(12)*,* 5547-5551.

Greenberg, R.A. (2005). Telomeres, crisis and cancer. *Current Molecular Medicine 5*(2), 213-218.

Greider, C.W. and Blackburn, E.H. (1987) The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell.* 51:887-898

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., & de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. *Cell, 97*(4), 503-514.

Guffei, A., Sarkar, R., Klewes, L., Righolt, C., Knecht, H., & Mai, S. (2010). Dynamic chromosomal rearrangements in Hodgkin's lymphoma are due to ongoing three-dimensional nuclear remodeling and breakage-bridge-fusion cycles. *Haematologica, 95*(12), 2038-2046. doi:10.3324/haematol.2010.030171

Hara, Y., Onishi, Y., Oishi, K., Miyazaki, K., Fukamizu, A., & Ishida, N. (2009). Molecular characterization of Mybbp1a as a co-repressor on the Period2 promoter. *Nucleic Acids Research, 37*(4), 1115-1126. doi:10.1093/nar/gkn1013

Harley, C.B., Futcher, A.B., & Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature, 345*(6274), 458-460.

Hau, M.P., & Tsao, S.W. (2017). Epstein-Barr virus hijacks DNA damage response transducers to orchestrate its life cycle. *Viruses, 9*(11), 341. doi:10.3390/v9110341

Hiyama, E., & Hiyama, K. (2007). Telomere and telomerase in stem cells. *British Journal of Cancer, 96*(7), 1020-1024. doi:10.1038/sj.bjc.6603671

Hoesel, B., & Schmidt, J. (2013). The complexity of NF-kB signalling in inflammation and cancer. *Molecular Cancer, 12,* 86. doi:10.1186/1476-4598-12-86

Holyz, H., Lipinska, N., Paszel-Jaworska, A., & Rubis, B. (2013). Telomerase as a useful target in cancer fighting--the breast cancer case. *Tumour Biology, 34*(3), 1371-1380. doi:10.1007/s13277-013-0757-4

Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker A., Gast A., … Kumar, R. (2013). TERT promoter mutations in familial and sporadic melanoma. *Science, 339*(6122), 959-961. doi:10.1126/science.1230062

Huen, D.S., Henderson, S.A., Croom-Carter, D. & Rowe, M. (1995). The Epstein-Barr virus latent membrane protein-1 (LMP1) mediates activation of NF-kappa B and cell surface phenotype via two effector regions in its carboxy-terminal cytoplasmic domain. *Oncogene, 10*(3), 549-560.

Jackson, S.P., & Bartek J. (2009). The DNA-damage response in human biology and disease. *Nature, 461*(7267), 1071-1078. doi:10.1038/nature08467

Jacobs, J. (2013). Loss of telomere protection: Consequences and opportunities. *Frontiers in Oncology 3,* 88. doi:10.3389/fonc.2013.00088

Jaffe, E.S., Harris, N.L., Stein, H., & Issacson, P. (2008). Classification of lymphoidneoplasms: the microscope as a tool for disease discovery. *Blood, 112*(12), 4384-4399.

Klein, U., & Dalla-Favera, R. (2008). Germinal centres: roles in B-cell physiology and malignancy. *Nature Reviews. Immunology. 8*(1), 22-23.

Knecht, H., Bachmann, E., Brousset, P., Sandvej, K., Nadal, D., Bachmann, F. … Pallesen, G. (1993). Deletions within the LMP1 oncogene of Epstein–Barr virus are clustered in Hodgkin's disease and identical to those observed in nasopharyngeal carcinoma. *Blood, 82*(10), 2937-2942.

Knecht, H., Berger, C., Rothenberger, S., Odermatt, B.F., & Brousset, P. (2001). The role of Epstein-Barr virus in neoplastic transformation. *Oncology. 60*(4), 289-302.

Knecht, H., Berger, C., McQuain, C., Rothenberger, S., Bachmann, E., Martin, J., … Odermatt, B.F. (1999). Latent membrane protein 1 associated signaling pathways are important in tumor cells of Epstein-Barr virus negative Hodgkin's disease. *Oncogene, 18*, 7161-7167.

Knecht, H., Sawan, B., Lichtensztejn, D., Lemieux, B., Wellinger R. J., & Mai, S. (2008). The 3D nuclear organization of telomeres marks the transition from Hodgkin to Reed-Sternberg cells. *Leukemia, 23*(3), 565-573. doi:10.1038/leu.2008.314

Knecht, H., Sawan, B., Lichtensztejn, Z., Lichtensztejn, D., & Mai, S. (2010). 3D Telomere FISH defines LMP1-expressing Reed–Sternberg cells as end-stage cells with telomere-poor 'ghost' nuclei and very short telomere. *Laboratory Investigation*, *90*(4), 611-619. doi:10.1038/labinvest.2010.2

Küppers, R., Rajewsky, K., Zhao, M., Simons, G., Laumann, R., Fischer, R., & Hansmann, M.L. (1994). Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. *Proceedings of the National Academy of Sciences USA, 91*(23), 10962-10966.

Küppers, R. (2009). The biology of Hodgkin's lymphoma. *Nature Reviews. Cancer,* 9(1), 15-27. doi:10.1038/nrc2542

Kutok, J.L., & Wang, F. (2006). Spectrum of Epstein-Barr virus-associated diseases. *Annual Review of Pathology, 1*, 375-404.

Lajoie, V., Lemieux, B., Sawan, B., Lichtensztejn, D., Lichtensztejn, Z., Wellinger, R., Knecht, H. (2015). LMP1 mediates multinuclearity through downregulation of shelterin proteins and formation of telomeric aggregates. *Blood*, *125*(13), 2101-2110.

LeBien, T.W., & Tedder, T. (2008). B lymphocytes: how they develop and function. *Blood,112*(5), 1570-1580. doi:10.1182/blood-2008-02-078071

Lee, I.S., Shin, Y.K., Chung, D.H., & Park, S.H. (2009). LMP1-induced downregulation of CD99 molecules in Hodgkin and Reed-Sternberg cells. *Leukemia & Lymphoma, 42*(4), 587-594. doi:10.3109/10428190109099318

Lee, M., & Blackburn, E. (1993). Sequence-specific DNA primer effects on telomerase polymerization activity. *Molecular and Cellular Biology, 13*(10), 6586-6599.

Leonard, S., & Wei, W., Anderton, J., Vockerodt, M., Rowe, M., Murray, P.G., & Woodman, C.B. (2011). Epigenetic and transcriptional changes which follow Epstein-Barr virus infection of germinal center B cells and their relevance to the pathogenesis of Hodgkin's lymphoma. *Journal of Virology, 85*(18), 9568-9577. doi:10.1128/JVI.00468-11

Li, L., Li, W., Xiao, L., Xu, J., Chen, X., Tang, M., … Cao, Y. (2012). Viral oncoprotein LMP1 disrupts p53-induced cell cycle arrest and apoptosis through modulating K63-linked ubiquitination of p53. *Cell Cycle, 11* (12), 2327-2336. doi:10.4161/cc.20771

Lowry, O.H., Rosenbrough, N.J, Farr, A.L. & Randall, R.J. (1951). Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry 163*(1), 265-275.

Maciejowski, J., & de Lange, T. (2017). Telomeres in cancer: tumor suppression and genome instability. *National Review of Molecular Cell Biology*, *18*(3), 175-186. doi:10.1038/nrm.2016.171

Mai, S., & Garini, Y. (2005). Oncogenic remodeling of the three-dimensional organization of the interphase nucleus: c-Myc induces telomeric aggregates whose formation precedes chromosomal rearrangements. *Cell Cycle. 4*(10), 1327-1331.

Mancao, C., & Hammerschmidt, W. (2007). Epstein-Barr virus latent membrane protein 2A is a B-cell receptor mimic and essential for B-cell survival. *Blood, 110*(10), 3715-3721. doi:10.1182/blood-2007-05-090142

Menezes, J., Leibold, W., Klein, G., & Clements, G. (1975). Establishment and characterization of an Epstein-Barr virus (EBC)-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma. *Biomedicine, 22*(4), 276-284.

Min, J. & Shay, J.W. (2016). TERT promoter mutations enhance telomerase activation by long-range chromatin interactions. *Cancer Discovery 6*(11), 1212-1214. doi:1158/2159-8290.CD-16-1050

Morris, M., Dawson, C., & Young, L. (2009). Role of the Epstein–Barr virus-encoded latent membrane protein-1, LMP1, in the pathogenesis of nasopharyngeal carcinoma. *Future Oncology 5*(6), 811-825. doi:10.2217/fon.09.53

Murnane, J. P. (2012). Telomere dysfunction and chromosome instability. *Mutation Research, 730*(1-2), 28-36. doi:10.1016/j.mrfmmm.2011.04.008

Nair, N., Shoaib, M., & Sorensen, C.S. (2017). Chromatin dynamics in genome stability: roles in suppressing endogenous DNA damage and facilitating DNA repair. *International Journal of Molecular Sciences, 18*(7), 1486. doi:10.3390/ijms18071486

New England Biolabs. (2015). Large (Klenow) Fragment Blunting (M0210). dx.doi.org/10.17504/protocols.io.cpgvjv

Oeseburg, H., de Boer, R.A., van Gilst, W.H., & van der Harst, P. (2009). Telomere biology in healthy aging and disease. *Pflugers Archiv – European Journal of Physiology, 459*(2), 259-268. doi:10/1007/s00424-009-0728-1

Palm, W., & de Lange, T. (2008). How shelterin protects mammalian telomeres. *Annual Review of Genetics, 42*, 301-334.

Phan, T.G., Paus, D., Chan, T.D., Turner, M.L., Nutt, S.L., Basten, A., & Brink, R. (2006). High affinity germinal center B cells are actively selected into the plasma cell compartment. *Journal of Experimental Medicine, 203*(11), 2419-2424. doi:10.1084/jem.20061254

Potter, H., & Heller, R. (2003). Transfection by electroporation. *Current Protocols in Molecular Biology,* 9(3). doi:10.1002/0471142727.mb0903s62

Rai, R., Chen, Y., Lei, M., & Chang, S. (2016). TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletion fusions. *Nature Communications, 7,* 10881. doi:10.1038/ncomms10881

Riley, K.J., Rabinowitz, G.S., Yario, T.A., Luna, J.M., Darnell, R.B., & Steitz, J.A. (2012). EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *EMBO Journal, 31*(9), 2207-2221. doi:0.1038/emboj.2012.63

Rengstl, B., Newrzela, S., Heinrich, T., Weiser, C., Thalheimer, Schmid, F., ... M.L. Hansmann. (2013). Incomplete cytokinesis and re-fusion of small mononucleated Hodgkin cells lead to giant multinucleated Reed- Sternberg cells. *Proceedings of the National Academy of Sciences USA, 110*(51), 20729-20734.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L., & Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature Genetics, 25*(3), 338-342.

Robertson K.D. (2001). DNA methylation, methyltransferases, and cancer. *Oncogene, 20*(24), 3139-3155.

Rountree, M.R., Bachman, K.E., & Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genetics, 25*(3), 269-277.

Schmutz, I., Timashev, L., Patel, J.D., & de Lange, T. (2017). TRF2 binds branched DNA to safeguard telomere integrity. *Nature Structural and Molecular Biology, 24*(9), 734-742. doi:10.1038/nsmb.3451

Shanbhag, S., & Ambinder, R.F. (2018). Hodgkin lymphoma: a review and update on recent progress. *CA: Cancer Journal for Clinicians, 68*(2), 116-132. doi:10.3322/caac.21438

Shippen-Lentz, D., & Blackburn, E.H. (1990). Functional evidence for an RNA template in telomerase. *Science, 247*(2942), 546-52. doi.org/10.1126/science.1689074

Soni, V., Cahir-McFarland, E., & Kieff, E. (2007). LMP1 trafficking activates growth and survival pathways. *Advances in Experimental Medicine and Biology, 597,* 173-187.

Stansel, R.M., de Lange, T., & Griffith, J.D. (2001). T-loop assembly in vitro involves binding of TRF2 near the 3′ telomeric overhang. *EMBO Journal, 20*(19), 5532-5540. doi:10.1093/emboi/20.19.5532

Stein, H. & Bob, R. (2009). Is Hodgkin lymphoma just another B-cell lymphoma? *Current Hematologic Malignancy Reports, 4*(3), 125-128. doi.org/10.1007/s11899-009-0018-1

Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., … Jaffe, E.S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood, 127*(20), 2375-2390. https://doi.org/10.1182/blood-2016-01-643569

Taylor, S. C., Laperriere, G., & Germain, H. (2017). Droplet digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Scientific Reports 7,* (2409).

ThermoFisher Scientific. (n.d.). *How a flow cytometer works.* Retrieved from https://www.thermofisher.com/us/en/home/life-science/cell-analysis/cell-analysis-learning-center/molecular-probes-school-of-fluorescence/flow-cytometry-basics/flow-cytometry-fundamentals/how-flow-cytometer-works.html

Thompson, C.A., Mauck, K., Havyer, R., Bhagra, A., Kalsi, H., & Hayes, S.N. (2011). Care of the adult Hodgkin lymphoma survivor. *American Journal of Medicine, 124*(12), 1106-1112. doi:10.1016/j.amjmed.2011.05.020

Thorley-Lawson, D.A., & Gross, A. (2004). Persistence of the Epstein-Barr virus and the origins of associated lymphomas. *New England Journal of Medicine, 350*(13), 1328-1337.

Tobon, G.J., Izquierdo, H. & Canas, C. (2013). B lymphocytes: Development, tolerance, and their role in autoimmunity-Focus on systemic Lupus erythematosus. *Autoimmune Diseases, 2013.* http://dx.doi.org/10.1155/2013/827254

Uchida, J., Yasui, T., Takaoka-Shichijo, Y., Muraoka, M., Kulwichit, W., Raab-Traub, N., & Kikutani H. (1999). Mimicry of CD40 signals by Epstein-Barr virus LMP1 in B lymphocyte responses. *Science, 286*(5438), 300-303.

Van den Hoff, M.J., Moorman, A.F., & Lamers, W.H. (1992). Electroporation in intracellular buffer increases cell survival. *Nucleic Acids Research, 20*(11), 2902.

Van Steensel, B., Smogorzewska, A. & de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. *Cell, 92*(3), 401-413.

Vaqué, J. P., Martínez, N., Batlle-López, A., Pérez, C., Montes-Moreno, S., Sánchez-Beato, M., & Piris, M. A. (2014). B-cell lymphoma mutations: improving diagnostics and enabling targeted therapies. *Haematologica, 99*(2), 222-231. doi:10.3324/haematol.2013.096248

Vaysberg, M., Lambert S.L., Krams, S.M., & Martinez, O.M. (2009) Activation of the JAK/STAT pathway in Epstein Barr virus+-associated posttransplant lymphoproliferative disease: role of interferon-gamma. *American Journal of Transplantation, 9*(10), 2292-2302. doi:10.1111/j.1600-6143.2009.02781.x

Vockerodt, M., Cader, F.Z., Shannon-Lowe, C. & Murray, P. (2014). Epstein-Barr virus and the origin of Hodgkin lymphoma. *Chinese Journal of Cancer, 33*(12), 591-597. doi:10.5732/cjc.014.10193

Vockerodt, M., Soares, M., Kanzler, H., Kuppers, R., Kube, D., Hansmann, M.L., … Tesch, H. (1998). Detection of clonal Hodgkin and Reed-Sternberg cells with identical somatically mutated and rearranged VH genes in different biopsies in relapsed Hodgkin's disease. *Blood. 92*(8), 2899-2907.

Watson, J.D. (1972). Origin of concatemeric T7 DNA. *Nature: New Biology, 239*(94), 197-201.

Weniger, M.A., & Kuppers, R. (2016). NF-kB deregulation in Hodgkin lymphoma. *Seminars in Cancer Biology, 39,* 32-39. doi:10.1016/j.semcancer.2016.05.001

Yakoob, J., Hu, G.L., Fan, X.-G. & Zhang, Z. (1999). Telomere, telomerase and digestive cancer. *World Journal of Gastroenterology, 5*(4), 334-337. doi:10.3748/wjg.v5.i4.334

Yao,Y., & Dai, W. (2014). Genomic instabilty and cancer. *Journal of Carcinogenesis and Mutagenesis, 5,* 110000165. doi:10.4172/2157-2518/1000165

Young, L.S., & Murray, P.G. (2003). Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene 22*(33), 5108-5121. doi:10.1038/sj.onc.1206556

Annex

RNA-Seq Data

Table RNA Seq Downregulated Genes

Gene	T-test	Q-value	Log2-ratio	Sample1 No LMP1	Sample2 LMP1	Sample2 No LMP1	Sample2 LMP1
JCHAIN	4.09E-02	2.36E-01	3.9	23.2	1.7	20.4	1.3
SPINK2	9.37E-02	2.55E-01	3.3	219.1	23.9	163.8	14.5
LTB.d	1.90E-02	2.25E-01	3.2	52.3	5.5	55.3	5.8
AIF1.b	5.83E-02	2.40E-01	3.0	13.8	1.5	11.7	1.7
KCNN3	7.64E-02	2.50E-01	3.0	11.5	1.3	9.3	1.2
LDLRAP1	1.46E-01	2.74E-01	2.9	14.8	3.4	23.2	1.7
AIF1.d	5.81E-02	2.40E-01	2.8	13.8	1.9	11.8	1.9
AIF1.c	5.81E-02	2.40E-01	2.8	13.8	1.9	11.8	1.9
AIF1	5.81E-02	2.40E-01	2.8	13.8	1.9	11.8	1.9
CA11	6.70E-02	2.47E-01	2.7	12.0	1.8	9.8	1.5
ITK	2.50E-02	2.25E-01	2.6	10.9	1.9	11.7	1.9
LRMP	8.69E-02	2.53E-01	2.6	198.2	30.7	157.3	28.9
DBN1	5.95E-02	2.40E-01	2.5	45.8	7.5	39.2	7.5
TRAF3IP3	3.01E-03	2.07E-01	2.4	12.3	2.5	12.1	2.1
MED20	4.38E-01	3.73E-01	2.4	26.3	3.5	5.3	2.7
PRKCE	4.03E-01	3.73E-01	2.3	11.4	3.0	44.7	8.4
RGS13	8.66E-03	2.11E-01	2.2	369.6	85.4	391.3	76.0
TNS3	2.56E-02	2.25E-01	2.1	18.9	2.6	17.1	5.6
NCAM2	1.54E-01	2.77E-01	2.1	15.9	1.0	9.4	4.8
RETREG1	1.47E-02	2.25E-01	2.1	116.9	28.3	108.0	24.5
MEIKIN	2.31E-01	3.09E-01	2.1	29.1	6.4	15.7	4.2
TRIB1	4.21E-02	2.36E-01	2.1	15.4	3.8	13.5	3.2
CD69	4.09E-02	2.36E-01	2.1	66.3	20.7	79.5	14.5
BCL6	1.04E-01	2.58E-01	2.0	16.7	4.2	12.4	2.9
NMU	8.49E-02	2.52E-01	1.9	40.8	9.2	33.2	10.1
NTNG1	5.97E-02	2.40E-01	1.8	11.4	3.1	10.0	3.0
LDLRAD4	7.08E-02	2.47E-01	1.8	12.6	3.0	10.4	3.5
SOX5	6.68E-03	2.08E-01	1.8	19.4	5.7	19.3	5.4
MAP3K7CL	1.54E-02	2.25E-01	1.8	17.9	5.7	18.1	4.7
UBL4A	5.39E-02	2.37E-01	1.8	14.0	6.3	18.1	3.0
DLG3	1.54E-02	2.25E-01	1.8	12.2	4.1	10.8	2.6
PTPN6	7.50E-02	2.49E-01	1.7	203.0	56.4	171.6	55.2
CAMK1	2.16E-03	1.95E-01	1.7	18.3	6.1	19.2	5.3
SMIM14	6.55E-02	2.45E-01	1.7	53.9	14.4	45.9	15.8
NUGGC	2.84E-01	3.37E-01	1.7	14.6	3.1	7.3	3.6
INPPL1	1.48E-01	2.74E-01	1.7	9.9	3.4	13.9	3.8
NT5DC2	1.30E-01	2.66E-01	1.7	21.1	5.7	15.8	5.6

SNX29	7.83E-03	2.08E-01	1.7	25.5	7.4	25.7	8.2
NME4	1.10E-01	2.59E-01	1.7	18.3	5.2	14.3	4.8
PXK	6.93E-02	2.47E-01	1.7	15.8	6.0	15.8	3.6
SYNJ2BP-COX16	4.90E-01	3.73E-01	1.7	3.7	2.8	18.9	4.1
ASMT	6.00E-02	2.40E-01	1.7	25.4	8.0	20.8	6.2
NLRP7	1.83E-01	2.92E-01	1.7	13.5	3.2	8.8	3.7
C1orf53	3.50E-01	3.59E-01	1.7	8.3	3.0	21.4	6.2
CCDC69	5.21E-03	2.08E-01	1.7	177.7	59.5	165.6	47.5
NUDT7	3.25E-01	3.51E-01	1.7	6.2	3.3	14.0	3.1
TSC22D1	2.13E-02	2.25E-01	1.7	13.7	4.1	15.4	5.0
GRAMD4	4.11E-01	3.73E-01	1.7	21.1	5.2	6.7	3.6
RALB	3.29E-01	3.53E-01	1.7	7.8	4.3	17.7	3.8
PITPNC1	1.10E-01	2.59E-01	1.7	22.8	6.6	18.0	6.4
HVCN1	1.10E-01	2.59E-01	1.6	57.7	16.7	45.4	16.3
NADK2	2.31E-02	2.25E-01	1.6	37.5	12.9	39.8	12.4
PLEKHA2	1.72E-02	2.25E-01	1.6	21.1	7.9	18.8	5.3
TMEM206	5.29E-01	3.85E-01	1.6	18.8	2.2	2.8	5.1
ECI2	7.13E-03	2.08E-01	1.6	20.1	6.9	20.1	6.6
MARCH6	1.37E-01	2.69E-01	1.6	20.7	8.2	27.8	8.3
LGALS9	2.71E-02	2.26E-01	1.6	38.5	13.8	41.2	13.3
CYFIP2	2.21E-02	2.25E-01	1.5	68.8	26.4	66.1	20.0
FCRLA	8.32E-03	2.09E-01	1.5	16.6	5.6	17.7	6.2
CORO1A	1.74E-01	2.87E-01	1.5	39.0	18.3	57.5	15.3
ZNF778	3.61E-01	3.65E-01	1.5	5.9	4.0	14.4	3.0
LMO2	1.72E-02	2.25E-01	1.5	51.9	18.1	48.8	17.1
PPP2R3B.b	5.88E-02	2.40E-01	1.5	16.7	5.6	14.8	5.6
TMEM116	6.06E-03	2.08E-01	1.5	17.1	6.6	18.1	6.0
RIOK1	1.15E-01	2.62E-01	1.5	21.0	8.1	27.0	9.1
DAAM1	1.01E-02	2.19E-01	1.5	14.4	5.2	14.1	5.2
MCTP2	5.93E-03	2.08E-01	1.4	14.5	5.6	13.5	4.8
CD72	3.59E-03	2.08E-01	1.4	16.7	6.4	16.0	5.8
PRKCB	3.76E-03	2.08E-01	1.4	58.9	21.0	61.6	24.3
POU2AF1	3.55E-02	2.33E-01	1.4	13.9	5.0	13.0	5.1
STX10	5.42E-02	2.37E-01	1.4	20.1	7.1	18.0	7.2
ERI3	3.44E-02	2.32E-01	1.4	13.2	3.9	11.2	5.3
PHF19	4.55E-02	2.37E-01	1.4	19.0	7.1	17.3	6.8
DOCK7	7.29E-02	2.48E-01	1.4	19.0	5.5	18.2	8.7
PTPN7	4.41E-03	2.08E-01	1.4	49.6	19.9	48.1	17.5
ALDH3A2	1.41E-02	2.24E-01	1.4	32.9	12.4	35.9	14.1
PTEN	3.89E-02	2.36E-01	1.4	703.2	264.6	644.5	253.8
MACC1	1.93E-02	2.25E-01	1.4	11.5	4.7	12.1	4.5
EVL	5.70E-02	2.39E-01	1.4	19.7	7.5	17.3	7.0
SDHAF4	2.72E-01	3.32E-01	1.3	19.7	7.1	37.8	15.4
SIGLEC10	6.59E-03	2.08E-01	1.3	22.4	9.1	21.0	8.0

TSTD1	2.31E-02	2.25E-01	1.3	30.5	13.1	29.8	11.0
PYCARD	1.97E-02	2.25E-01	1.3	21.8	8.3	22.1	9.4
TUBA4A	1.49E-02	2.25E-01	1.3	32.9	14.7	30.6	11.4
APPL2	3.10E-02	2.30E-01	1.3	10.9	4.4	10.0	4.2
RALGPS2	2.10E-02	2.25E-01	1.3	20.6	9.6	21.5	7.7
HDAC9	6.77E-02	2.47E-01	1.3	60.4	26.3	68.7	27.1
ADARB1	1.12E-01	2.60E-01	1.3	12.8	3.4	9.3	5.8
SMIM11B	2.08E-01	2.96E-01	1.3	20.7	13.3	22.0	4.6
LCP2	3.05E-02	2.29E-01	1.2	24.5	10.6	26.2	10.9
ITPR2	8.78E-05	1.24E-01	1.2	31.2	13.4	30.9	13.2
DVL3	1.06E-01	2.58E-01	1.2	52.4	34.0	71.3	19.0
GPX7	5.49E-03	2.08E-01	1.2	14.9	6.2	15.1	6.7
ZNF236	4.09E-02	2.36E-01	1.2	13.7	6.0	15.2	6.4
CHM	6.34E-01	4.20E-01	1.2	4.4	24.8	65.9	5.8
ADAM15	1.67E-01	2.85E-01	1.2	23.2	6.6	24.0	14.0
CD3G	6.96E-02	2.47E-01	1.2	15.0	6.2	13.3	6.2
EEPD1	4.22E-02	2.36E-01	1.2	14.3	7.1	12.5	4.7
VNN2	5.75E-02	2.39E-01	1.2	17.5	9.1	20.6	7.8
PLEKHF2	1.71E-03	1.88E-01	1.2	35.8	16.7	37.0	15.5
RFXANK	1.01E-01	2.58E-01	1.2	15.5	5.5	16.0	8.5
LIMD2	2.66E-02	2.26E-01	1.2	16.6	6.9	16.8	7.9
EIF2B3	3.06E-01	3.42E-01	1.2	162.4	17.6	121.9	109.0
PKN2	4.87E-01	3.73E-01	1.2	9.9	8.3	32.5	10.6
SNX22	5.60E-03	2.08E-01	1.2	12.7	5.6	12.5	5.6
ATP8A1	8.84E-02	2.53E-01	1.2	129.8	50.9	108.7	55.5
ARV1	2.37E-02	2.25E-01	1.1	14.2	5.9	13.1	6.4
AC022016.2	5.61E-02	2.38E-01	1.1	34.4	19.3	37.4	13.2
DTX1	8.16E-02	2.50E-01	1.1	11.7	4.3	11.3	6.2
ETV5	1.38E-02	2.24E-01	1.1	11.3	4.6	10.7	5.5
STAMBPL1	2.00E-02	2.25E-01	1.1	25.9	11.7	25.0	11.7
MZB1	3.27E-02	2.32E-01	1.1	51.9	24.0	57.3	26.2
AC134669.2	1.10E-01	2.59E-01	1.1	13.3	4.7	10.2	6.1
GSN	1.65E-02	2.25E-01	1.1	11.0	5.0	10.5	4.9
DCXR	4.75E-02	2.37E-01	1.1	240.3	113.9	203.9	92.7
RASGRP3	3.75E-03	2.08E-01	1.1	92.2	41.6	95.9	46.0
RNASET2	8.01E-02	2.50E-01	1.1	13.2	6.3	10.6	4.8
TERF2	3.92E-01	3.73E-01	1.1	30.3	10.8	13.7	9.7
TMEM256	3.47E-02	2.32E-01	1.1	19.2	9.9	17.1	7.1
TMEM256.b	3.47E-02	2.32E-01	1.1	19.2	9.9	17.1	7.1
PAGR1	1.68E-03	1.88E-01	1.1	17.9	8.2	18.2	8.7
DOK3	1.41E-01	2.71E-01	1.1	10.6	5.4	13.8	6.2
CD22	1.82E-02	2.25E-01	1.1	46.0	23.4	49.4	21.7
PNKD	1.39E-02	2.24E-01	1.1	13.7	6.1	14.8	7.4
CD180	4.97E-02	2.37E-01	1.1	12.4	6.0	13.8	6.5
ZC3H12D	1.07E-01	2.58E-01	1.1	10.5	5.4	12.5	5.5

TK2	1.75E-01	2.88E-01	1.1	12.8	3.3	8.9	7.2
HSBP1L1	4.71E-02	2.37E-01	1.1	13.6	6.4	12.5	6.2
STIM2	6.71E-02	2.47E-01	1.1	39.4	18.2	35.2	17.8
STRBP	1.30E-01	2.66E-01	1.0	96.6	42.0	77.9	42.5
CD79A	8.36E-03	2.09E-01	1.0	36.6	18.4	35.3	16.4
TKFC	3.48E-01	3.59E-01	1.0	15.5	5.3	8.0	6.2
PRDX2	1.96E-02	2.25E-01	1.0	55.6	27.2	51.8	25.4
ZDHHC23	3.51E-01	3.59E-01	1.0	7.0	7.7	14.0	2.6
TEC	3.38E-02	2.32E-01	1.0	10.7	5.2	10.0	5.0
ODF2L	1.55E-01	2.77E-01	1.0	27.1	18.6	30.3	9.8
SNX25	3.93E-02	2.36E-01	1.0	91.8	48.3	91.8	42.5
BPNT1	1.06E-01	2.58E-01	1.0	46.5	21.8	39.1	21.1
MSI2	3.69E-02	2.33E-01	1.0	24.5	13.0	24.2	11.4
DPP3	1.32E-01	2.66E-01	1.0	26.3	10.2	19.9	13.0
SGSM3	7.06E-02	2.47E-01	1.0	34.1	15.0	33.7	19.0
ZCCHC2	4.01E-01	3.73E-01	1.0	11.4	8.9	24.1	9.1
CPNE5	1.53E-02	2.25E-01	1.0	11.5	5.8	11.2	5.8
S100A4	1.47E-01	2.74E-01	1.0	55.6	33.7	71.4	30.7
SMYD3	1.82E-02	2.25E-01	1.0	16.7	8.7	15.2	7.6
PLA2G12A	8.44E-02	2.52E-01	1.0	21.5	10.5	18.6	10.0
CDCA7L	4.21E-02	2.36E-01	1.0	53.0	28.7	45.6	21.7
CDK14	2.70E-02	2.26E-01	1.0	38.8	19.5	37.1	19.3
BLOC1S2	5.24E-01	3.82E-01	1.0	234.7	61.7	72.0	95.6
LMCD1	1.17E-01	2.63E-01	1.0	9.9	5.7	11.9	5.6
NT5C3B	1.80E-01	2.91E-01	1.0	35.6	16.3	26.8	16.0
USP22	3.73E-01	3.71E-01	1.0	35.6	27.1	69.2	27.1
PIK3AP1	5.42E-02	2.37E-01	0.9	151.3	72.6	136.3	76.6
HMGA1	7.62E-02	2.49E-01	0.9	91.4	45.7	79.9	43.1
DIS3L	4.55E-01	3.73E-01	0.9	10.5	11.6	26.0	7.4
AC270107.10	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
AC270107.5	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
AC270107.4	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
AC270107.7	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
AC270107.9	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
AC270107.3	8.10E-02	2.50E-01	0.9	22.1	10.1	18.8	11.2
EEF1A2	9.36E-02	2.55E-01	0.9	47.4	23.6	40.8	22.6
CARHSP1	1.02E-01	2.58E-01	0.9	11.5	5.4	9.8	5.8
HADH	1.36E-01	2.68E-01	0.9	12.8	8.2	13.1	5.4
AKT1	1.28E-01	2.65E-01	0.9	24.3	12.0	20.0	11.5
ARHGAP25	7.70E-03	2.08E-01	0.9	24.0	12.0	22.8	12.8
SUGCT	2.91E-01	3.37E-01	0.9	16.2	8.7	9.5	4.9
LMO4	1.36E-01	2.68E-01	0.9	17.2	7.8	13.7	8.6
SUPT4H1	2.94E-02	2.28E-01	0.9	59.5	33.0	59.4	30.4
RBM41	3.95E-01	3.73E-01	0.9	16.8	6.4	33.4	20.3
BMP2K	4.95E-02	2.37E-01	0.9	23.6	12.9	25.5	13.2

ACADM	3.99E-02	2.36E-01	0.9	55.8	31.0	59.3	30.6
LRP11	1.53E-01	2.76E-01	0.9	14.0	8.5	10.6	4.7
MXD4	1.53E-01	2.76E-01	0.9	12.7	6.1	10.1	6.1
IKZF1	1.00E-03	1.87E-01	0.9	26.6	13.9	26.1	14.4
ODC1	1.52E-02	2.25E-01	0.9	345.0	171.0	332.6	193.3
CBFB	1.50E-01	2.75E-01	0.9	74.5	35.5	59.4	36.6
ATE1	5.41E-01	3.88E-01	0.9	20.3	20.9	65.4	25.2
CLECL1	1.53E-01	2.76E-01	0.9	73.6	45.5	92.5	44.1
PDE7A	4.40E-04	1.85E-01	0.9	20.4	10.9	20.6	11.2
CCNY	1.96E-01	2.95E-01	0.9	63.5	29.1	47.1	30.5
CHN2	1.99E-02	2.25E-01	0.9	21.1	10.3	19.3	11.5
TMC6	1.32E-01	2.66E-01	0.9	19.1	9.8	15.2	8.7
CD79B	3.96E-02	2.36E-01	0.9	191.0	96.1	174.3	102.4
NUDT1	5.12E-02	2.37E-01	0.9	14.1	8.0	15.2	7.9
KANK1	1.33E-01	2.66E-01	0.9	14.3	8.2	18.1	9.5
DCK	2.98E-02	2.28E-01	0.9	36.4	20.7	36.1	19.0
NPM3	4.33E-02	2.37E-01	0.9	74.9	36.0	65.7	41.5
PEBP1	7.06E-02	2.47E-01	0.9	53.5	27.0	47.4	28.6
CEP250	2.93E-01	3.37E-01	0.9	15.5	12.0	24.8	10.3
SDHAF3	1.48E-01	2.74E-01	0.8	40.0	22.1	30.0	16.7
FCHO1	7.78E-02	2.50E-01	0.8	16.2	8.8	13.9	7.9
APEH	6.85E-02	2.47E-01	0.8	61.1	32.3	55.4	32.6
GAK	4.97E-02	2.37E-01	0.8	28.3	16.3	30.3	16.4
ASPH	4.62E-02	2.37E-01	0.8	43.1	26.5	44.5	22.3
CAMSAP1	9.89E-02	2.58E-01	0.8	44.8	24.7	36.6	20.8
METTL26	3.27E-02	2.32E-01	0.8	14.9	8.3	16.6	9.3
LYSMD2	4.75E-02	2.37E-01	0.8	15.1	9.0	15.0	7.9
LYRM1	2.45E-02	2.25E-01	0.8	50.0	29.1	50.0	27.4
RMI2	3.08E-02	2.30E-01	0.8	12.5	7.3	11.2	6.1
NFS1	2.11E-02	2.25E-01	0.8	14.1	8.3	14.9	8.0
PRPS2	7.64E-02	2.50E-01	0.8	29.9	16.3	26.5	15.6
CDV3	4.97E-03	2.08E-01	0.8	205.7	118.6	198.1	109.6
DPM3	8.32E-03	2.09E-01	0.8	32.0	19.5	33.5	17.6
AC068547.1	2.46E-02	2.25E-01	0.8	15.9	9.3	15.9	8.8
CUX1	8.29E-02	2.51E-01	0.8	44.2	26.3	50.0	27.3
SLC27A5	6.49E-02	2.45E-01	0.8	12.0	5.5	10.1	7.1
MYO1E	1.48E-02	2.25E-01	0.8	120.4	69.6	111.6	62.7
HGS	1.59E-01	2.78E-01	0.8	58.4	28.8	46.2	30.9
PSD4	1.23E-01	2.65E-01	0.8	9.0	6.9	11.5	4.8
SERPINB6	3.39E-02	2.32E-01	0.8	15.0	9.0	13.6	7.3
GCDH	1.88E-03	1.92E-01	0.8	40.0	22.8	39.8	22.8
NPIPB13	9.57E-03	2.19E-01	0.8	11.1	6.3	10.8	6.2
AK3	1.23E-03	1.87E-01	0.8	11.2	6.6	11.5	6.4
SEL1L3	2.40E-02	2.25E-01	0.8	12.3	7.3	11.3	6.2
CHEK2	3.19E-02	2.31E-01	0.8	15.5	8.7	17.4	10.1

GON7.b	1.54E-02	2.25E-01	0.8	12.8	7.5	13.2	7.4
GON7	1.54E-02	2.25E-01	0.8	12.8	7.5	13.2	7.4
MAP4K1	1.81E-02	2.25E-01	0.8	25.6	15.1	25.1	13.9
UBE2J1	1.36E-02	2.23E-01	0.8	10.3	6.5	11.1	5.8
DHRS4L2	3.63E-01	3.66E-01	0.8	15.3	10.6	9.6	3.6
MAIP1	4.60E-01	3.73E-01	0.8	10.1	8.9	22.3	9.7
PLEKHO1	3.36E-02	2.32E-01	0.8	16.7	9.4	15.9	9.3
STAU2	1.12E-01	2.59E-01	0.8	41.8	27.6	50.2	25.1
SMARCAL1	5.24E-01	3.82E-01	0.8	18.9	18.6	51.1	21.6
CDK5	9.23E-02	2.54E-01	0.8	11.1	6.2	13.7	8.1
TIMM8B	1.46E-02	2.25E-01	0.8	42.1	23.7	45.4	26.6
ANXA4	1.36E-01	2.68E-01	0.8	13.9	9.3	14.0	6.7
CCDC138	1.01E-01	2.58E-01	0.8	40.4	24.8	46.4	25.2
SKAP1	2.63E-02	2.26E-01	0.8	14.4	8.3	13.1	7.5
BIK	7.75E-02	2.50E-01	0.8	13.4	6.8	11.5	7.6
CNN3	3.63E-02	2.33E-01	0.8	52.0	31.6	46.8	25.6
NIPSNAP1	7.39E-03	2.08E-01	0.8	19.0	11.2	18.0	10.3
CHD7	7.58E-02	2.49E-01	0.8	11.0	6.6	12.7	7.2
LRRTM4	3.47E-02	2.32E-01	0.8	11.4	6.0	10.3	6.6
TMC8	3.69E-02	2.33E-01	0.8	11.6	7.1	12.4	6.9
SNRPN	6.83E-03	2.08E-01	0.8	33.3	18.4	32.3	19.9
IFT81	1.13E-01	2.60E-01	0.8	19.6	13.0	19.7	10.0
GUK1	1.91E-01	2.95E-01	0.8	322.0	156.7	243.7	174.0
WDR54	8.54E-02	2.52E-01	0.8	15.6	8.9	13.3	8.0
HMOX2.b	2.46E-01	3.18E-01	0.8	21.2	10.8	15.1	10.4
NPIPB4	6.97E-02	2.47E-01	0.8	75.2	42.4	89.5	54.1
AIG1	5.32E-03	2.08E-01	0.8	12.7	7.6	12.3	7.1
MZT2B	1.82E-02	2.25E-01	0.8	45.1	25.1	44.2	27.3
ITGB2	6.31E-02	2.44E-01	0.8	36.6	24.4	41.9	21.8
CBR4	7.95E-02	2.50E-01	0.8	25.9	14.9	22.4	13.6
PRKRA	1.22E-02	2.23E-01	0.8	23.3	14.4	24.4	13.8
CDK2AP1	2.44E-02	2.25E-01	0.8	45.3	26.4	42.5	25.3
ARPC4-TTLL3	1.82E-01	2.92E-01	0.8	18.9	14.8	21.3	8.9
PRPSAP2	2.41E-02	2.25E-01	0.8	81.6	49.6	74.5	42.5
OCIAD2	4.33E-02	2.37E-01	0.8	62.4	42.7	68.2	34.4
AC073610.3	7.10E-02	2.48E-01	0.8	18.4	9.8	16.2	10.7
ANAPC15	1.17E-01	2.63E-01	0.8	42.3	26.5	50.1	28.2
EVI2A	1.02E-01	2.58E-01	0.8	16.5	10.2	19.3	11.0
PTPRC.b	8.06E-03	2.08E-01	0.8	108.9	68.1	114.4	64.2
COMT	3.05E-01	3.42E-01	0.8	19.1	18.6	25.6	7.9
E2F5	1.16E-01	2.63E-01	0.8	19.3	10.9	16.3	10.2
KLF13	1.95E-01	2.95E-01	0.8	29.7	17.9	21.3	12.4
ATP5G1	3.22E-01	3.50E-01	0.7	36.3	17.2	23.0	18.1
C1D	3.51E-02	2.33E-01	0.7	35.3	21.5	36.9	21.6

TTC7B	3.22E-01	3.50E-01	0.7	10.6	6.5	17.5	10.3
SLAIN1	2.96E-02	2.28E-01	0.7	10.3	6.2	11.2	6.7
COTL1	1.23E-01	2.65E-01	0.7	44.8	24.7	38.2	25.1
TRIM27.g	3.27E-01	3.52E-01	0.7	11.2	4.3	12.8	10.0
PAQR3	8.12E-02	2.50E-01	0.7	12.4	7.6	14.3	8.4
MMADHC	5.38E-01	3.87E-01	0.7	74.5	73.9	196.4	89.2
GPI.b	1.64E-01	2.82E-01	0.7	26.4	15.9	19.9	12.0
H3F3A	2.86E-02	2.28E-01	0.7	249.0	142.7	233.9	148.4
PPP6R1	3.33E-02	2.32E-01	0.7	40.3	22.8	39.5	25.3
AES	1.09E-02	2.21E-01	0.7	59.1	35.5	62.6	38.0
RAB4A	2.10E-02	2.25E-01	0.7	20.9	11.6	19.3	12.7
CCDC91	1.28E-01	2.65E-01	0.7	195.2	129.6	229.8	127.1
MAPK10	1.15E-02	2.23E-01	0.7	30.5	18.4	28.9	17.4
TTC39C	6.47E-02	2.45E-01	0.7	11.1	8.0	13.0	6.6
AKT2	3.42E-01	3.59E-01	0.7	17.3	12.6	28.3	15.1
PTPRC	2.36E-02	2.25E-01	0.7	35.0	23.7	38.3	20.8
AL132780.3	2.58E-01	3.24E-01	0.7	10.1	7.9	10.2	4.5
MYCBP	4.75E-02	2.37E-01	0.7	29.3	16.1	28.1	18.9
CYB5R4	3.24E-02	2.32E-01	0.7	29.3	18.4	29.1	17.2
BCL7B	1.76E-02	2.25E-01	0.7	14.4	9.1	14.0	8.3
TRIM27.c	3.31E-01	3.54E-01	0.7	11.2	4.5	12.9	10.2
AFMID	2.71E-02	2.26E-01	0.7	25.6	15.4	28.2	17.4
NDUFB10	4.91E-02	2.37E-01	0.7	111.6	63.5	102.0	66.9
GALNT14	5.64E-02	2.38E-01	0.7	14.7	9.5	14.7	8.5
MGST3	4.45E-01	3.73E-01	0.7	18.7	9.9	9.3	7.3
GATAD2A	6.81E-03	2.08E-01	0.7	25.6	16.6	26.6	15.4
KLHL8	1.53E-01	2.76E-01	0.7	16.6	11.5	20.2	11.1
CDC40	4.20E-01	3.73E-01	0.7	25.4	12.5	13.7	11.5
HIST2H2AB	4.86E-02	2.37E-01	0.7	780.0	426.0	692.1	477.0
BIN2	6.27E-03	2.08E-01	0.7	18.5	11.8	19.3	11.3
POLE3	7.69E-02	2.50E-01	0.7	94.7	59.4	109.0	65.6
RNF180	5.00E-02	2.37E-01	0.7	23.5	13.7	23.6	15.2
MRPL42	5.41E-02	2.37E-01	0.7	95.1	57.5	86.0	54.0
SRA1	3.57E-01	3.63E-01	0.7	19.8	10.1	12.1	9.6
FAM216A	2.66E-02	2.26E-01	0.7	47.3	28.9	44.2	27.5
PRKD3	8.95E-02	2.53E-01	0.7	37.0	22.4	31.5	19.8
TXNIP	2.67E-03	2.04E-01	0.7	10.2	6.3	10.0	6.1
ACTR6	4.80E-01	3.73E-01	0.7	36.0	41.5	77.8	28.9
NDUFAF8	8.01E-03	2.08E-01	0.7	213.6	124.4	205.2	134.7
SON	6.74E-01	4.33E-01	0.7	32.2	77.7	167.1	45.6
TRPT1	3.26E-01	3.52E-01	0.7	32.7	18.5	20.5	14.5
DTD2	1.14E-02	2.23E-01	0.7	13.4	8.6	13.9	8.3
BAG1	1.17E-01	2.63E-01	0.7	66.3	38.8	57.4	38.0
UCHL1	3.23E-02	2.32E-01	0.7	47.5	28.0	46.9	30.6
ULK4	4.95E-02	2.37E-01	0.7	24.9	16.4	23.9	14.0

CEP192	6.01E-02	2.40E-01	0.7	10.6	6.8	11.5	7.0
HIST1H4H	2.29E-01	3.09E-01	0.7	56.2	43.0	75.9	39.2
TFB1M	3.29E-02	2.32E-01	0.7	21.2	12.7	23.4	15.0
NEK6	4.33E-02	2.37E-01	0.7	17.1	9.5	15.3	10.7
COCH	4.94E-03	2.08E-01	0.7	36.9	23.2	35.5	21.9
CSK	1.56E-02	2.25E-01	0.7	44.2	27.9	44.1	27.1
MAPK9	1.21E-01	2.65E-01	0.7	13.2	8.7	15.8	9.4
MYC	2.78E-02	2.28E-01	0.7	38.2	26.4	40.7	22.8
PET100	1.38E-01	2.69E-01	0.7	434.3	273.0	341.4	210.7
ELL3	1.81E-01	2.91E-01	0.7	25.5	14.2	20.5	14.5
VPS13A	3.64E-01	3.67E-01	0.7	19.3	6.0	13.3	14.3
MEF2C	3.97E-02	2.36E-01	0.7	170.3	116.2	187.4	107.2
MAZ	4.86E-03	2.08E-01	0.7	41.9	25.1	40.4	26.3
KIFC1	1.47E-01	2.74E-01	0.7	9.8	7.5	10.7	5.3
RAP1A	8.64E-03	2.11E-01	0.7	30.6	19.9	31.8	19.2
NCF1	2.26E-01	3.07E-01	0.7	22.6	12.3	17.1	12.5
GRAMD2A	4.88E-01	3.73E-01	0.7	9.9	7.3	21.4	12.3
SSBP4	1.75E-02	2.25E-01	0.7	11.5	7.4	10.9	6.6
PEX1	2.07E-01	2.96E-01	0.7	12.6	10.4	17.1	8.2
DCUN1D4	5.33E-02	2.37E-01	0.7	17.0	9.1	14.9	10.9
AP000275.2	7.96E-02	2.50E-01	0.7	93.6	56.3	85.2	56.0
COX5A	1.09E-01	2.59E-01	0.7	54.1	32.6	46.6	30.7
MRPL20	2.90E-01	3.37E-01	0.7	205.1	104.9	141.1	112.7
SERPINA9	1.72E-02	2.25E-01	0.7	27.2	16.6	28.1	18.2
ITGB1BP1	1.61E-01	2.80E-01	0.7	40.8	23.8	33.4	22.8
AK9	9.46E-02	2.55E-01	0.7	9.7	6.9	11.3	6.3
EPHA7	1.65E-02	2.25E-01	0.7	19.9	13.5	20.8	12.2
POLD4	1.63E-01	2.81E-01	0.7	20.3	14.6	19.2	10.3
SFXN4	1.61E-02	2.25E-01	0.7	15.9	9.9	15.4	9.8
CYTIP	6.51E-02	2.45E-01	0.7	14.0	9.1	15.3	9.4
TMCO3	6.92E-02	2.47E-01	0.7	11.2	8.5	12.8	6.7
ROCK2	3.38E-02	2.32E-01	0.7	33.3	20.7	31.7	20.4
CD9	9.64E-02	2.56E-01	0.7	12.3	7.8	10.2	6.4
GOT2	9.84E-02	2.58E-01	0.7	29.3	17.4	26.1	17.5
SGK1	5.21E-02	2.37E-01	0.7	75.4	54.7	85.8	47.2
NECAP2	8.21E-02	2.51E-01	0.7	49.4	29.5	44.7	30.0
GSR	8.37E-04	1.85E-01	0.7	22.7	14.1	22.5	14.5
PARP1	1.19E-02	2.23E-01	0.7	74.6	48.5	75.4	46.3
TRIAP1	2.85E-02	2.28E-01	0.7	26.5	16.2	25.3	16.6
CDCA7	2.94E-02	2.28E-01	0.7	21.1	13.6	21.8	13.6
RAPGEF5	1.91E-01	2.95E-01	0.7	20.0	10.6	22.4	16.4
PTK2	2.04E-01	2.95E-01	0.7	25.2	16.2	34.3	21.7
IMPDH1	5.95E-02	2.40E-01	0.7	37.3	22.5	38.0	25.5
BOLA2B	1.38E-01	2.69E-01	0.6	107.2	76.4	104.6	58.7
PEX16	3.01E-01	3.40E-01	0.6	21.2	9.8	22.0	17.9

SYNGR2	3.82E-02	2.33E-01	0.6	82.7	49.6	76.6	52.1
C12orf45	1.66E-01	2.84E-01	0.6	37.1	20.0	29.5	22.6
RPGR	3.70E-01	3.70E-01	0.6	10.5	7.5	17.7	10.5
MSL3	8.31E-02	2.51E-01	0.6	44.8	29.9	49.3	30.2
OSBPL8	7.82E-02	2.50E-01	0.6	41.3	25.4	37.8	25.2
MDM2	1.10E-01	2.59E-01	0.6	104.0	73.3	104.5	60.2
ACSF3	1.93E-01	2.95E-01	0.6	11.5	6.8	9.0	6.3
TMEM154	1.89E-01	2.95E-01	0.6	17.1	12.7	15.8	8.5
ANAPC5	8.02E-03	2.08E-01	0.6	50.5	32.0	51.1	33.3
TRIM28	1.31E-01	2.66E-01	0.6	87.6	44.9	78.9	62.2
CERCAM	4.78E-01	3.73E-01	0.6	834.9	616.8	1703.2	1017.4
WDR20	3.38E-02	2.32E-01	0.6	11.2	7.3	12.0	7.7
PRDX6	2.10E-03	1.92E-01	0.6	36.5	23.0	35.9	23.8
GNAS	5.33E-03	2.08E-01	0.6	103.0	67.5	100.4	64.0
CARS2	2.62E-01	3.27E-01	0.6	9.0	6.0	13.0	8.3
UBAC1	3.93E-02	2.36E-01	0.6	74.2	43.6	70.6	50.2
ERP29	5.67E-03	2.08E-01	0.6	92.4	59.9	89.4	57.9
AC138969.1	3.85E-01	3.73E-01	0.6	46.1	21.6	27.6	26.2
GGA2	1.18E-02	2.23E-01	0.6	27.3	17.7	26.0	16.8
RERE	1.83E-01	2.92E-01	0.6	11.1	7.9	13.7	8.1
DONSON	1.02E-01	2.58E-01	0.6	33.2	22.6	37.3	23.2
HDAC8	7.01E-01	4.41E-01	0.6	6.2	18.6	31.8	6.1
CFAP36	1.91E-02	2.25E-01	0.6	25.9	16.7	24.5	16.0
CIT	4.40E-01	3.73E-01	0.6	7.3	4.5	13.2	8.8
NSMAF	3.42E-01	3.59E-01	0.6	26.9	19.2	42.6	26.0
ZC3H14	1.68E-01	2.85E-01	0.6	38.4	21.5	31.0	23.6
MPP6	2.86E-02	2.28E-01	0.6	31.8	19.3	29.5	20.6
HERC3	2.48E-01	3.19E-01	0.6	15.3	9.7	21.5	14.3
VAV3	1.20E-01	2.65E-01	0.6	11.8	9.1	12.8	6.9
CMC4	5.13E-02	2.37E-01	0.6	10.9	7.0	10.1	6.7
HDAC1	3.62E-03	2.08E-01	0.6	76.5	49.8	75.4	49.3
TPGS2	1.77E-02	2.25E-01	0.6	39.2	25.8	36.6	23.7
CAMK2D	7.49E-02	2.49E-01	0.6	10.8	6.2	9.4	7.0
STUB1	7.72E-02	2.50E-01	0.6	164.4	111.7	179.2	112.9
RHOBTB3	1.92E-01	2.95E-01	0.6	127.1	83.7	96.2	62.5
ATPIF1	5.25E-02	2.37E-01	0.6	376.8	246.1	335.8	220.9
SMIM4	1.26E-01	2.65E-01	0.6	28.9	23.1	35.4	19.1
PSMD5	1.46E-01	2.74E-01	0.6	10.0	6.7	12.6	8.2
MZT2A	1.45E-01	2.73E-01	0.6	80.4	50.5	66.4	45.9
PDLIM7	4.95E-02	2.37E-01	0.6	13.3	10.0	14.6	8.4
TERF1	2.53E-02	2.25E-01	0.6	12.0	7.5	11.3	7.8
GLUD1	1.61E-02	2.25E-01	0.6	92.8	65.9	98.9	60.2
CLUAP1	2.10E-01	2.97E-01	0.6	12.0	6.9	9.5	7.2
POP5	5.26E-02	2.37E-01	0.6	24.1	15.2	24.3	16.7
NLK	2.22E-02	2.25E-01	0.6	10.8	7.2	10.0	6.6

NUCB2	1.97E-01	2.95E-01	0.6	343.2	255.2	299.1	168.0
ABRACL	4.80E-02	2.37E-01	0.6	231.3	158.5	231.4	146.5
TMEM156	2.95E-02	2.28E-01	0.6	24.9	18.2	27.2	16.2
UBE2H	3.09E-02	2.30E-01	0.6	16.2	9.8	14.8	10.7
MAP4K1.b	2.83E-02	2.28E-01	0.6	27.5	18.1	25.4	16.8
TCTN1	7.74E-02	2.50E-01	0.6	12.5	7.7	13.0	9.2
TPD52	1.48E-02	2.25E-01	0.6	65.9	46.8	70.1	43.1
OXLD1	1.82E-01	2.92E-01	0.6	14.5	8.9	11.7	8.4
COX8A	1.05E-01	2.58E-01	0.6	23.6	14.3	23.5	16.9
OIP5	2.88E-01	3.37E-01	0.6	10.2	7.3	14.8	9.3
CD53	2.21E-02	2.25E-01	0.6	89.6	55.2	83.5	59.4
OXNAD1	8.09E-02	2.50E-01	0.6	10.6	6.8	9.7	6.7
CERS2	3.14E-02	2.31E-01	0.6	20.4	13.5	18.7	12.4
PPP1R10	7.90E-01	4.69E-01	0.6	1.4	4.2	47.8	28.5
IRAK1	7.85E-02	2.50E-01	0.6	12.5	9.0	14.0	8.5
PPIP5K2	1.08E-01	2.58E-01	0.6	29.8	21.2	33.6	20.8
SPICE1	6.06E-02	2.41E-01	0.6	18.5	12.9	16.8	10.6
KCTD19	1.94E-01	2.95E-01	0.6	16.0	11.2	20.7	13.2
CRLS1	2.22E-01	3.04E-01	0.6	10.4	7.5	13.5	8.4
ZSWIM7	9.91E-02	2.58E-01	0.6	133.4	84.5	119.9	83.6
PTDSS1	1.28E-02	2.23E-01	0.6	50.2	31.2	47.5	33.7
KMT5A	3.70E-02	2.33E-01	0.6	55.9	37.7	58.3	38.1
STK24	7.57E-02	2.49E-01	0.6	40.7	26.0	37.6	26.1
NECAB3	3.28E-01	3.53E-01	0.6	12.0	5.4	11.0	9.9
ARPC1B	7.47E-02	2.49E-01	0.6	101.6	66.1	90.6	61.8
RNASEH2B	2.64E-01	3.27E-01	0.6	645.7	348.5	470.1	394.3
INVS	8.44E-02	2.52E-01	0.6	11.0	7.0	10.1	7.0
MSANTD3	7.80E-03	2.08E-01	0.6	18.4	12.5	17.8	11.7
SYK	3.58E-02	2.33E-01	0.6	36.6	23.7	36.6	25.1
GAPT	1.44E-01	2.73E-01	0.6	46.1	33.8	54.1	32.9
EEF1B2	9.73E-02	2.56E-01	0.6	862.6	507.4	742.5	562.5
EEF1B2.b	9.73E-02	2.56E-01	0.6	862.6	507.4	742.5	562.5
FGD6	7.48E-02	2.49E-01	0.6	12.0	7.7	11.0	7.6
STIM1	5.77E-02	2.39E-01	0.6	18.5	12.2	16.7	11.3
PHF10	3.77E-02	2.33E-01	0.6	15.6	10.1	14.8	10.2
POLR2G	1.79E-02	2.25E-01	0.6	26.3	16.3	24.6	17.7
SYCP1	7.07E-02	2.47E-01	0.6	12.4	9.1	12.8	7.8
SEPT6	1.20E-02	2.23E-01	0.6	60.7	40.7	57.5	38.1
GCSH	8.22E-02	2.51E-01	0.6	77.9	48.4	70.1	50.5
CENPM	5.61E-02	2.38E-01	0.6	13.0	7.5	11.8	9.1
ABHD10	5.13E-03	2.08E-01	0.6	43.1	28.8	44.3	29.6
TOP2B	3.16E-02	2.31E-01	0.6	46.1	31.5	47.7	31.3
MRPL36	4.14E-02	2.36E-01	0.6	148.1	95.0	139.5	97.5
USP11	1.41E-01	2.71E-01	0.6	11.6	7.0	9.9	7.4
CLK3	2.02E-01	2.95E-01	0.6	14.7	9.2	11.6	8.5

POLR3GL	5.22E-02	2.37E-01	0.6	18.9	13.1	17.0	11.0
ENOPH1	1.75E-02	2.25E-01	0.6	30.3	21.5	32.1	20.3
MPC2	1.17E-01	2.63E-01	0.6	33.8	20.2	33.1	24.6
SPC24	1.12E-01	2.59E-01	0.6	24.3	14.6	21.0	15.7
TMEM230	1.67E-01	2.85E-01	0.6	32.1	25.0	39.3	22.9
ALAS1	9.33E-02	2.55E-01	0.6	14.3	9.5	12.2	8.3
ARPC2	2.46E-02	2.25E-01	0.6	471.3	323.5	467.1	307.1
LBR	5.81E-02	2.40E-01	0.6	26.6	18.6	26.2	16.9
MTCP1	1.27E-01	2.65E-01	0.6	31.9	20.3	27.9	19.9
MTAP	2.64E-01	3.27E-01	0.6	172.7	101.4	129.2	101.8
STN1	9.76E-02	2.57E-01	0.6	16.1	10.9	13.7	9.1
TMEM141	4.14E-02	2.36E-01	0.6	19.5	13.5	20.4	13.4
SLC4A7	6.32E-02	2.44E-01	0.6	44.8	29.3	42.0	29.2
ENAH	2.09E-02	2.25E-01	0.6	21.0	14.3	19.6	13.1
SLFN11	6.26E-02	2.43E-01	0.6	25.7	17.5	22.6	15.0
CENPV	2.42E-02	2.25E-01	0.6	107.0	73.8	100.1	66.0
FYN	1.06E-01	2.58E-01	0.6	59.8	38.7	53.2	37.7
PPAT	2.52E-02	2.25E-01	0.6	31.7	21.3	34.3	23.3
PHC1	1.31E-01	2.66E-01	0.6	12.0	8.3	14.6	9.7
SMN2.c	9.87E-02	2.58E-01	0.6	26.3	17.0	23.8	16.9
POC1B	2.96E-01	3.37E-01	0.6	19.3	10.3	21.6	17.4
LPCAT3	1.35E-01	2.68E-01	0.6	13.2	11.2	15.6	8.3
PLPBP	8.70E-02	2.53E-01	0.6	21.7	12.5	18.6	14.8
AARSD1	2.89E-02	2.28E-01	0.6	24.4	15.0	22.4	16.7
TMEM14B	7.89E-02	2.50E-01	0.6	43.4	31.2	43.5	27.7
AMZ2	7.23E-03	2.08E-01	0.6	42.6	28.9	41.0	27.7
BMPR1A	2.19E-01	3.03E-01	0.6	20.2	11.4	15.6	12.9
ATP6V1G1	3.61E-02	2.33E-01	0.6	139.0	97.6	131.2	85.7
CENPW	1.45E-01	2.73E-01	0.6	62.8	40.8	52.5	37.4
PET117	2.21E-02	2.25E-01	0.6	11.6	7.8	11.1	7.6
POLR3K	2.13E-02	2.25E-01	0.6	19.0	12.2	18.5	13.2
NFE2L3	3.90E-02	2.36E-01	0.6	15.1	10.2	13.8	9.4
PUS7	1.33E-02	2.23E-01	0.6	21.9	15.2	21.4	14.3
TIMMDC1	8.58E-02	2.52E-01	0.6	35.2	23.0	32.2	22.8
SYPL1	2.15E-02	2.25E-01	0.6	10.8	7.3	10.2	7.0
ARL6IP4	1.32E-01	2.66E-01	0.6	205.0	130.0	179.0	131.2
MTFR1	1.53E-01	2.76E-01	0.6	33.8	25.4	39.8	24.6
CTBP1	1.36E-02	2.23E-01	0.6	14.8	10.7	15.7	10.0
TIAM2	8.01E-02	2.50E-01	0.6	11.5	8.4	11.7	7.4
MRPS27	4.28E-02	2.37E-01	0.6	122.2	85.8	121.6	80.4
GRPEL1	1.52E-02	2.25E-01	0.6	37.2	26.3	38.8	25.5
NDUFB11	1.42E-01	2.72E-01	0.6	16.0	9.6	13.4	10.5
LTA4H	3.10E-02	2.30E-01	0.6	51.2	31.7	47.6	35.8
FAM3C	4.12E-02	2.36E-01	0.6	20.9	15.4	21.8	13.7
TCF4	1.31E-01	2.66E-01	0.6	66.0	49.8	76.5	47.5

ACO2	3.92E-02	2.36E-01	0.6	14.3	9.4	14.8	10.5
MAGEF1	3.10E-02	2.30E-01	0.6	20.1	13.8	18.5	12.6

Table 14: Downregulated genes in the presence of LMP1-induction found via RNA Seq

Table RNA Seq Upregulated Genes

Gene	T-test	Q-value	Log2-ratio	Sample1 No LMP1	Sample2 LMP1	Sample2 No LMP1	Sample2 LMP1
DAPP1	2.79E-03	2.06E-01	-0.6	21.1	30.9	21.6	31.6
IFT46	1.77E-01	2.89E-01	-0.6	9.3	15.5	12.5	16.4
LYN	7.92E-03	2.08E-01	-0.6	20.6	31.3	21.3	30.2
MYH9	1.45E-01	2.74E-01	-0.6	27.7	40.4	33.7	49.7
NFKBIE	1.12E-01	2.59E-01	-0.6	17.9	27.3	17.1	24.0
SLC38A2	7.05E-02	2.47E-01	-0.6	25.1	43.2	30.5	38.4
DNAJA1	1.84E-02	2.25E-01	-0.6	226.5	341.0	247.0	354.6
CALM3	3.00E-02	2.28E-01	-0.6	275.7	395.5	263.1	396.7
B2M	3.55E-02	2.33E-01	-0.6	660.6	999.5	667.5	959.2
NRP2	2.02E-02	2.25E-01	-0.6	9.2	13.9	9.9	14.3
BAZ2B	2.57E-02	2.25E-01	-0.6	12.2	18.3	12.2	17.8
ZNF682	1.07E-01	2.58E-01	-0.6	8.1	11.1	6.6	10.7
CD86	6.24E-02	2.43E-01	-0.6	77.9	120.9	85.8	121.5
SLC5A6	7.01E-02	2.47E-01	-0.6	10.3	16.1	11.7	16.5
TM9SF4	1.30E-01	2.66E-01	-0.6	6.5	10.6	7.9	10.7
TYW3	3.98E-02	2.36E-01	-0.6	28.0	43.3	32.3	46.3
PRKAG1	6.17E-02	2.41E-01	-0.6	14.9	23.3	18.0	25.5
SCYL2	1.63E-01	2.81E-01	-0.6	6.9	13.0	9.7	11.7
UBE2B	3.72E-02	2.33E-01	-0.6	17.7	27.4	19.7	28.3
TLR1	1.95E-01	2.95E-01	-0.6	5.7	10.4	8.0	9.9
CTSH	4.16E-02	2.36E-01	-0.6	51.0	69.2	46.6	76.2
USP9X	4.65E-02	2.37E-01	-0.6	11.2	18.1	12.6	17.5
SQLE	1.66E-01	2.83E-01	-0.6	175.7	284.7	175.1	238.7
MPZ	2.54E-01	3.23E-01	-0.6	9.8	12.8	5.9	10.5
APOBEC3G	1.05E-02	2.21E-01	-0.6	11.2	17.3	11.5	16.7
QPRT	1.58E-01	2.77E-01	-0.6	12.2	19.9	12.4	16.8
PLOD1	8.26E-02	2.51E-01	-0.6	9.4	15.6	11.1	15.1
DBI	5.08E-03	2.08E-01	-0.6	453.9	682.8	446.7	668.7
RAB12	1.47E-01	2.74E-01	-0.6	6.4	11.1	7.1	9.2
ACSL1	1.68E-02	2.25E-01	-0.6	64.9	102.8	68.2	97.4
NUP62	1.01E-01	2.58E-01	-0.6	16.7	28.9	20.7	27.4
TFG	1.51E-02	2.25E-01	-0.6	16.8	24.1	15.7	24.7
POC5	2.33E-01	3.10E-01	-0.6	10.4	19.6	15.5	19.3
SFT2D1	2.61E-02	2.26E-01	-0.6	53.2	86.1	57.0	80.0
NPC1	1.15E-01	2.62E-01	-0.6	14.8	25.3	16.2	21.6
ISG15	2.18E-01	3.03E-01	-0.6	19.9	26.8	20.5	34.2

HLA-DQB1.e	1.07E-01	2.58E-01	-0.6	10.2	18.4	12.1	15.3
RIDA	5.19E-02	2.37E-01	-0.6	43.9	67.6	42.2	62.5
RAB3IP	3.53E-03	2.08E-01	-0.6	21.1	31.8	20.2	30.8
TAGAP	8.52E-02	2.52E-01	-0.6	8.5	14.1	9.0	12.5
GOLGB1	1.96E-02	2.25E-01	-0.6	9.5	14.4	10.2	15.4
FKTN	4.95E-01	3.73E-01	-0.6	59.0	80.1	101.4	163.2
CD55	9.30E-02	2.54E-01	-0.6	6.9	11.0	6.8	9.9
ABHD3	9.24E-02	2.54E-01	-0.6	31.4	58.8	40.5	50.7
IRF2BP2	1.55E-01	2.77E-01	-0.6	8.0	14.0	10.4	14.0
KIAA1191	6.35E-02	2.44E-01	-0.6	16.7	27.2	17.3	24.9
ZNF33A	2.76E-02	2.27E-01	-0.6	17.5	26.8	16.1	24.7
SEC61B	8.22E-02	2.51E-01	-0.6	55.8	82.3	56.8	90.5
ANXA5	9.67E-02	2.56E-01	-0.6	51.0	82.0	50.6	73.7
STAT3	2.73E-02	2.26E-01	-0.6	8.8	14.6	9.5	13.4
ABHD2	2.85E-02	2.28E-01	-0.6	7.9	12.5	8.5	12.7
ARID2	5.85E-01	4.05E-01	-0.6	8.0	6.0	6.4	16.2
IFI30	4.58E-02	2.37E-01	-0.6	87.7	120.8	78.6	134.9
PLSCR1	6.69E-02	2.47E-01	-0.6	15.7	25.3	15.9	23.4
RAB8B	5.26E-02	2.37E-01	-0.6	41.6	70.9	45.4	63.6
ARAP1	5.70E-01	3.98E-01	-0.6	12.2	9.5	9.9	24.6
EPB41	3.32E-02	2.32E-01	-0.6	44.7	68.3	46.4	72.8
SMARCA2	2.48E-02	2.25E-01	-0.6	29.5	46.7	30.8	46.8
ZNF140	1.43E-01	2.72E-01	-0.6	8.9	15.1	9.1	12.8
DNAJB4	8.73E-02	2.53E-01	-0.6	12.4	21.0	14.5	20.8
MLLT6.b	2.66E-01	3.28E-01	-0.6	5.5	9.7	9.4	13.5
SASH3	2.45E-01	3.17E-01	-0.6	14.2	20.2	16.6	27.9
ST20	1.35E-01	2.68E-01	-0.7	13.2	23.6	18.5	26.2
KRAS	5.05E-02	2.37E-01	-0.7	6.6	11.6	7.8	11.1
HLA-B.f	2.42E-02	2.25E-01	-0.7	91.3	139.8	85.3	138.0
ACSL5	4.37E-02	2.37E-01	-0.7	42.2	65.2	43.2	69.4
CNTNAP4	1.99E-02	2.25E-01	-0.7	6.5	10.6	7.1	10.9
IKBKE	5.48E-02	2.37E-01	-0.7	12.6	21.3	15.7	23.7
MBD5	1.43E-01	2.73E-01	-0.7	8.4	12.6	9.3	15.5
HELLS	1.75E-01	2.88E-01	-0.7	189.0	394.3	247.4	298.6
LAP3	1.61E-02	2.25E-01	-0.7	82.2	142.5	91.2	133.1
CCDC124	6.73E-02	2.47E-01	-0.7	25.7	42.2	25.3	38.8
UBE2L6	2.10E-02	2.25E-01	-0.7	18.4	29.2	19.7	31.3
B4GALT4	1.49E-01	2.74E-01	-0.7	8.3	15.0	8.8	12.3
RAB21	3.53E-02	2.33E-01	-0.7	42.7	78.0	50.5	70.8
BAZ2A	2.01E-01	2.95E-01	-0.7	6.9	9.6	6.8	12.3
ARHGAP45	1.13E-02	2.23E-01	-0.7	6.2	10.4	6.5	9.9
EIF2AK2	1.81E-01	2.91E-01	-0.7	17.9	25.9	18.5	32.3
TNIP2	3.54E-02	2.33E-01	-0.7	28.4	49.0	31.9	47.6
MT-ND6	2.26E-01	3.07E-01	-0.7	63.8	134.4	101.4	131.1

FYTTD1	5.65E-02	2.38E-01	-0.7	77.6	131.6	86.8	132.6
CLPTM1	7.65E-03	2.08E-01	-0.7	9.1	14.3	8.4	13.7
HLA-B.c	6.48E-03	2.08E-01	-0.7	10.2	15.6	9.7	16.3
PPP3CC	4.39E-02	2.37E-01	-0.7	7.0	12.1	7.9	11.8
FAHD2A	9.85E-04	1.87E-01	-0.7	8.4	13.4	8.3	13.4
LRTOMT	6.03E-01	4.11E-01	-0.7	12.4	18.1	2.8	6.5
SVIP	7.47E-02	2.49E-01	-0.7	6.6	10.9	6.3	9.9
RHOQ	1.40E-03	1.87E-01	-0.7	38.5	61.1	37.6	61.6
STK40	5.60E-02	2.38E-01	-0.7	12.0	20.9	13.7	20.5
SPPL2A	9.16E-02	2.53E-01	-0.7	20.5	29.6	16.7	30.4
RGS5	4.39E-02	2.37E-01	-0.7	8.8	16.6	10.6	14.8
HIST1H2BM	1.98E-01	2.95E-01	-0.7	12.8	19.1	14.5	25.1
HSPA1B.b	7.83E-02	2.50E-01	-0.7	28.0	52.7	35.2	49.6
HLA-A	1.45E-01	2.74E-01	-0.7	38.8	68.0	38.2	56.8
AP1S3	2.19E-02	2.25E-01	-0.7	14.0	22.9	15.5	24.9
CCDC93	1.59E-01	2.78E-01	-0.7	11.7	17.7	12.8	22.0
SLC25A12	5.80E-01	4.02E-01	-0.7	16.2	13.2	16.1	39.2
PMAIP1	9.36E-02	2.55E-01	-0.7	19.4	34.6	20.5	30.2
N4BP1	8.15E-02	2.50E-01	-0.7	7.4	13.0	9.9	15.1
LACTB	5.75E-02	2.39E-01	-0.7	12.1	17.8	10.2	18.4
ARL8B	1.64E-01	2.82E-01	-0.7	37.4	54.4	37.0	66.8
CBX4	7.23E-03	2.08E-01	-0.7	8.1	13.6	8.7	14.0
C2orf76	7.05E-02	2.47E-01	-0.7	22.4	36.3	18.7	30.9
METTL1	1.01E-01	2.58E-01	-0.7	6.0	11.6	7.0	9.6
ZNF12	5.04E-01	3.74E-01	-0.7	3.9	19.1	18.1	16.9
PPP1R9B	1.49E-02	2.25E-01	-0.7	18.6	30.2	18.9	31.3
MTRNR2L12	3.41E-01	3.59E-01	-0.7	18.9	26.7	24.2	44.0
HLA-B	7.87E-02	2.50E-01	-0.7	79.6	120.8	76.4	135.2
OPTN	3.34E-02	2.32E-01	-0.7	43.5	72.3	41.4	67.2
DTX3L	2.35E-02	2.25E-01	-0.7	12.1	20.1	13.6	22.0
YIPF3	3.93E-02	2.36E-01	-0.7	17.7	30.8	21.6	34.1
IKBKB	1.89E-01	2.95E-01	-0.7	4.7	10.7	7.6	9.6
SLC3A2	2.40E-02	2.25E-01	-0.7	37.3	67.6	42.5	64.6
ANKRD12	3.31E-02	2.32E-01	-0.7	149.3	280.5	177.4	261.9
HTRA2	5.32E-02	2.37E-01	-0.7	9.1	14.9	9.8	16.6
PHACTR1	2.34E-02	2.25E-01	-0.7	5.6	9.7	6.5	10.3
RBM19	5.34E-02	2.37E-01	-0.7	8.3	13.9	9.5	15.8
STX3	1.93E-02	2.25E-01	-0.7	11.9	19.2	10.3	17.9
ARNTL2	4.29E-02	2.37E-01	-0.7	11.9	20.8	13.0	20.8
TANK	4.08E-02	2.36E-01	-0.8	102.4	183.2	113.8	180.8
LPCAT1	7.96E-02	2.50E-01	-0.8	23.9	43.7	25.1	38.8
DNAJB9	7.86E-02	2.50E-01	-0.8	45.2	66.9	35.9	70.1
PSD3	6.10E-02	2.41E-01	-0.8	25.2	46.6	32.2	50.4
HLA-B.d	1.93E-02	2.25E-01	-0.8	57.4	103.5	60.7	97.0
CD58	3.53E-03	2.08E-01	-0.8	58.2	96.9	56.9	99.1

FGFR1	5.99E-02	2.40E-01	-0.8	6.6	10.8	6.6	11.7
PIAS1	2.33E-02	2.25E-01	-0.8	10.6	18.2	11.5	19.7
PREX1	7.91E-02	2.50E-01	-0.8	25.8	40.5	19.3	37.3
CHCHD6	1.31E-01	2.66E-01	-0.8	11.0	16.7	8.0	16.3
LRR1	2.65E-02	2.26E-01	-0.8	294.6	537.5	350.4	577.6
MT-ATP6	6.77E-02	2.47E-01	-0.8	145.2	270.3	194.3	317.3
FBXW11	4.31E-01	3.73E-01	-0.8	14.9	34.0	14.4	16.8
PSMB9.h	4.88E-02	2.37E-01	-0.8	12.8	23.5	14.6	23.9
NPC2	2.83E-02	2.28E-01	-0.8	5.6	9.6	6.0	10.5
XBP1	2.58E-02	2.25E-01	-0.8	26.5	50.3	30.2	48.6
DHRS7B	1.94E-02	2.25E-01	-0.8	6.1	10.2	5.7	10.3
PTTG1IP	8.45E-02	2.52E-01	-0.8	15.1	23.5	13.9	27.2
SQSTM1	9.44E-02	2.55E-01	-0.8	17.3	28.9	18.2	33.3
DPH6	1.37E-01	2.69E-01	-0.8	16.4	24.6	11.7	24.6
SPOCK2	2.56E-02	2.25E-01	-0.8	9.9	19.8	11.4	17.9
CD70	4.83E-02	2.37E-01	-0.8	45.5	74.2	39.2	75.8
EIF1	1.85E-02	2.25E-01	-0.8	21.5	39.1	24.1	42.6
RBM7	2.23E-01	3.05E-01	-0.8	5.2	14.0	9.9	13.1
CTSS	2.87E-02	2.28E-01	-0.9	63.0	112.2	65.2	119.2
TRIM27.e	3.73E-01	3.71E-01	-0.9	13.3	29.7	12.0	16.1
PSME2	4.39E-03	2.08E-01	-0.9	212.0	384.9	206.1	374.2
SERPINB9	2.04E-02	2.25E-01	-0.9	5.1	10.6	6.1	9.9
UNC119	3.31E-02	2.32E-01	-0.9	15.5	25.3	13.0	26.6
HLA-C.g	1.32E-02	2.23E-01	-0.9	49.7	94.9	51.9	90.3
NR6A1	1.52E-02	2.25E-01	-0.9	8.0	14.8	8.9	16.0
FTH1	1.29E-01	2.65E-01	-0.9	90.6	153.4	95.1	186.5
HSPB1	6.51E-02	2.45E-01	-0.9	6.9	13.8	8.3	14.0
STARD4	1.87E-01	2.94E-01	-0.9	18.4	38.0	17.7	28.6
LITAF	2.50E-02	2.25E-01	-0.9	10.9	20.0	9.8	18.1
HLA-B.b	3.54E-02	2.33E-01	-0.9	63.1	139.4	78.7	122.4
PARP9	1.60E-02	2.25E-01	-0.9	19.9	40.2	22.7	38.6
RHOG	8.20E-03	2.08E-01	-0.9	10.3	19.0	10.6	19.6
POTEI	3.66E-02	2.33E-01	-0.9	5.9	10.4	5.3	10.4
CD274	1.53E-01	2.76E-01	-0.9	8.5	19.8	13.0	20.0
ACTB	1.18E-03	1.87E-01	-0.9	776.1	1438.1	761.1	1414.6
HLA-E.b	2.15E-01	3.00E-01	-0.9	17.6	31.7	10.8	21.1
OAS1	1.84E-01	2.92E-01	-0.9	11.9	18.9	11.7	25.0
STAT2	6.24E-02	2.43E-01	-0.9	7.7	14.0	8.4	16.0
SPIB	2.17E-02	2.25E-01	-0.9	14.9	27.8	16.1	30.1
UGCG	2.72E-05	4.50E-02	-0.9	9.4	17.5	9.4	17.6
MKNK2	3.58E-02	2.33E-01	-0.9	13.8	27.2	15.4	27.4
MX1	5.68E-02	2.39E-01	-0.9	35.9	54.8	28.2	65.2
VRK2	1.45E-02	2.25E-01	-0.9	14.3	30.1	16.7	28.1
DOCK10	1.17E-01	2.63E-01	-0.9	8.0	18.0	11.6	18.9
PCMTD2	6.97E-02	2.47E-01	-0.9	6.4	13.7	7.1	11.8

CERS4	5.42E-03	2.08E-01	-0.9	6.8	12.2	6.5	12.7
IFT52	5.91E-01	4.06E-01	-0.9	7.5	5.3	7.7	23.4
CDC42SE1	4.06E-02	2.36E-01	-0.9	12.4	22.3	12.0	24.0
IFI16	1.38E-02	2.24E-01	-0.9	17.1	35.4	19.5	33.8
IFIH1	1.68E-02	2.25E-01	-0.9	10.4	20.3	11.1	20.4
HLA-C.f	4.44E-03	2.08E-01	-0.9	24.5	49.6	26.7	47.5
ISG20	1.04E-02	2.21E-01	-0.9	11.1	22.5	12.2	21.9
ZNF260	1.34E-01	2.67E-01	-0.9	87.6	210.6	131.1	205.4
RAPGEF1	1.32E-03	1.87E-01	-0.9	11.0	21.2	11.5	21.8
IL10	6.80E-02	2.47E-01	-0.9	19.2	42.9	24.7	41.3
CRYZ	8.26E-02	2.51E-01	-0.9	14.1	26.1	15.4	30.6
LEPROT	1.26E-01	2.65E-01	-0.9	4.9	14.3	8.5	11.3
OAS2	7.90E-02	2.50E-01	-0.9	5.7	10.5	6.1	12.1
SNORD3B-1	6.97E-02	2.47E-01	-0.9	52.2	99.0	59.6	116.3
PELI1	7.42E-02	2.49E-01	-0.9	8.1	16.4	7.9	14.6
JUNB	2.02E-02	2.25E-01	-1.0	8.5	17.0	9.1	17.1
SLF1	3.63E-01	3.67E-01	-1.0	25.0	64.3	25.7	33.9
STX12	6.13E-02	2.41E-01	-1.0	8.8	16.8	9.9	19.4
LSAMP	1.15E-01	2.62E-01	-1.0	18.1	29.5	12.7	30.2
TNNT1	7.97E-02	2.50E-01	-1.0	28.9	54.9	33.0	65.4
ZCCHC18	5.97E-01	4.08E-01	-1.0	0.3	16.2	10.0	3.9
CCL5.b	3.06E-02	2.29E-01	-1.0	9.8	19.0	10.6	20.8
FLOT2	1.75E-01	2.88E-01	-1.0	15.8	24.0	9.0	24.6
SLC6A6	9.90E-02	2.58E-01	-1.0	8.4	17.3	7.8	14.5
MT-CYB	8.50E-02	2.52E-01	-1.0	55.6	125.1	73.5	127.8
LTBR	8.22E-02	2.51E-01	-1.0	19.4	39.2	25.3	48.9
TRIM27	3.51E-01	3.59E-01	-1.0	11.3	29.4	11.6	15.8
JDP2	3.72E-02	2.33E-01	-1.0	9.8	17.9	7.8	17.0
C1orf54	4.94E-02	2.37E-01	-1.0	8.5	16.4	8.9	18.1
DUSP2	1.81E-02	2.25E-01	-1.0	16.9	30.2	14.3	31.7
CYBB	1.63E-02	2.25E-01	-1.0	7.7	15.3	7.8	15.7
TSPAN33	7.34E-03	2.08E-01	-1.0	11.8	25.8	13.2	24.2
IFI35	1.09E-01	2.59E-01	-1.0	5.8	10.9	6.2	13.2
IFITM1	2.27E-02	2.25E-01	-1.0	18.1	36.8	17.8	35.2
CASP1	4.64E-03	2.08E-01	-1.0	196.2	401.9	214.3	422.4
CCL5	7.45E-02	2.49E-01	-1.0	8.9	21.5	13.6	24.2
SUOX	1.13E-01	2.60E-01	-1.0	6.1	10.3	4.2	10.4
PDXK	1.45E-03	1.87E-01	-1.0	23.0	46.8	23.3	47.4
ARID3A	1.18E-02	2.23E-01	-1.0	10.3	22.2	12.2	24.0
RNF213	2.89E-01	3.37E-01	-1.0	6.0	10.0	7.2	17.0
FAM98B	6.21E-01	4.16E-01	-1.0	2.1	35.9	17.9	5.2
NCK2	1.26E-01	2.65E-01	-1.0	9.9	24.2	11.1	19.2
IRF8	4.69E-02	2.37E-01	-1.1	99.0	186.7	76.9	178.2
NDUFB4	4.72E-01	3.73E-01	-1.1	25.8	25.5	21.1	72.2
SEC14L1	1.70E-01	2.85E-01	-1.1	12.0	23.3	14.6	32.3

TRPS1	1.06E-02	2.21E-01	-1.1	5.4	11.0	5.1	11.0
LAT2	3.65E-02	2.33E-01	-1.1	24.1	53.8	25.0	49.9
MAP3K4	4.23E-01	3.73E-01	-1.1	17.6	28.0	28.4	69.5
BST2	1.52E-01	2.75E-01	-1.1	138.9	502.9	297.1	424.9
IFNGR2	4.47E-02	2.37E-01	-1.1	15.6	34.2	15.2	31.5
CD82	2.86E-02	2.28E-01	-1.1	4.8	11.3	6.4	12.9
EXTL2	8.32E-02	2.51E-01	-1.1	9.1	16.1	7.6	19.8
PRKCZ	6.99E-02	2.47E-01	-1.1	12.9	33.4	17.6	32.3
MT2A	1.54E-01	2.76E-01	-1.1	6.7	13.8	8.2	18.8
RAB11FIP1	1.72E-02	2.25E-01	-1.1	9.3	19.3	8.3	18.9
ABCC4	6.59E-03	2.08E-01	-1.1	24.8	58.6	27.4	55.2
CASP4	4.72E-03	2.08E-01	-1.1	31.6	70.1	32.3	70.2
TFRC	5.97E-02	2.40E-01	-1.1	73.4	170.9	74.5	154.0
IVNS1ABP	1.72E-02	2.25E-01	-1.1	42.1	97.2	43.6	92.4
MXI1	1.03E-01	2.58E-01	-1.1	16.1	38.0	15.4	31.6
ALCAM	7.95E-03	2.08E-01	-1.2	11.8	26.6	12.8	28.2
NCKAP1	3.45E-01	3.59E-01	-1.2	7.7	113.8	74.6	70.9
SLC39A1	1.75E-03	1.88E-01	-1.2	9.3	20.2	8.7	20.7
AGPAT3	2.45E-02	2.25E-01	-1.2	8.3	16.7	7.2	18.6
VPS26A	2.28E-02	2.25E-01	-1.2	14.0	37.2	17.6	35.3
FSD1L	3.18E-02	2.31E-01	-1.2	4.3	10.8	5.9	12.6
LONRF1	4.48E-02	2.37E-01	-1.2	10.2	27.2	11.8	23.7
FCHSD2	3.64E-03	2.08E-01	-1.2	70.2	162.5	68.6	158.9
SKIL	3.09E-02	2.30E-01	-1.2	5.6	12.6	5.6	13.4
MT-CO1	1.92E-02	2.25E-01	-1.2	281.9	697.9	357.3	785.0
CTSB	8.08E-03	2.08E-01	-1.2	8.7	19.9	7.9	18.6
SZT2	5.68E-01	3.98E-01	-1.2	3.8	4.8	10.7	29.1
HLA-C.c	6.72E-02	2.47E-01	-1.2	5.4	13.4	5.4	11.9
VCAM1	1.88E-02	2.25E-01	-1.2	10.4	26.7	13.3	28.6
MT-ND2	2.64E-02	2.26E-01	-1.2	107.4	278.5	147.3	318.9
NABP1	4.80E-02	2.37E-01	-1.2	26.0	68.6	32.7	69.7
CD84	7.56E-02	2.49E-01	-1.2	6.2	14.3	7.0	17.0
ATP1B1	2.12E-02	2.25E-01	-1.2	9.0	23.7	10.7	23.1
TRIM69.b	5.73E-02	2.39E-01	-1.3	7.0	14.7	5.0	14.1
RNF145	3.82E-03	2.08E-01	-1.3	12.3	32.4	13.8	30.7
LRRC32	2.33E-02	2.25E-01	-1.3	4.6	10.5	4.1	10.5
CSRP2	5.39E-02	2.37E-01	-1.3	8.7	21.2	7.5	18.2
HLA-E.g	1.74E-01	2.87E-01	-1.3	13.1	25.4	11.8	35.8
STAT5A	4.40E-03	2.08E-01	-1.3	8.7	21.0	7.8	19.8
SLC2A5	5.03E-02	2.37E-01	-1.3	5.3	13.7	5.1	12.4
DENND5A	4.72E-02	2.37E-01	-1.3	4.1	10.9	4.2	9.9
CCDC28B	3.55E-02	2.33E-01	-1.3	71.9	184.0	84.9	210.6
C1orf61	9.66E-02	2.56E-01	-1.3	4.0	11.8	7.0	16.0
NFKB2	8.83E-04	1.85E-01	-1.3	31.1	82.8	33.2	80.9
CLIP2	3.90E-02	2.36E-01	-1.4	7.8	21.1	8.0	19.4

CSTB	1.18E-03	1.87E-01	-1.4	4.8	12.6	5.2	13.0
DUSP22	1.38E-01	2.70E-01	-1.4	64.8	196.6	70.0	148.8
IFI6	1.94E-01	2.95E-01	-1.4	6.6	13.1	6.1	19.4
MT-ND4	1.64E-02	2.25E-01	-1.4	107.5	308.1	145.2	342.4
FN3KRP	1.65E-01	2.83E-01	-1.4	13.8	28.2	4.4	18.7
CD80	1.30E-02	2.23E-01	-1.4	8.9	25.4	11.3	27.0
WDFY1	2.16E-01	3.01E-01	-1.4	16.1	26.1	10.5	43.2
MT-ND5	4.60E-02	2.37E-01	-1.4	70.7	221.1	102.1	232.6
SSTR2	4.26E-02	2.37E-01	-1.4	14.3	37.4	16.1	42.4
FMNL3	3.60E-02	2.33E-01	-1.4	5.2	15.1	6.2	15.0
IFI44	3.74E-02	2.33E-01	-1.4	10.3	29.4	13.7	34.9
PPP1R15A	6.93E-02	2.47E-01	-1.4	4.8	13.7	4.7	11.9
NAMPT	1.48E-01	2.74E-01	-1.4	73.4	251.3	87.3	181.5
RAB9A	5.67E-02	2.39E-01	-1.4	30.3	87.9	31.3	78.2
PPP4R4	6.50E-04	1.85E-01	-1.4	5.2	14.6	5.5	14.3
HERPUD1	2.23E-02	2.25E-01	-1.4	8.9	27.6	10.4	24.7
NEDD4L	4.06E-02	2.36E-01	-1.5	5.0	14.9	6.9	17.9
RGS1	4.23E-02	2.36E-01	-1.5	88.3	312.5	120.2	259.9
CFLAR	9.89E-03	2.19E-01	-1.5	25.7	72.8	29.5	78.9
XAF1	4.10E-02	2.36E-01	-1.5	3.0	9.5	4.6	11.6
DENND4A	8.01E-02	2.50E-01	-1.5	20.3	51.0	19.8	60.0
TCFL5	1.74E-02	2.25E-01	-1.5	4.1	11.5	4.5	12.3
HYOU1	6.91E-02	2.47E-01	-1.5	18.1	53.0	16.8	45.4
HYOU1.b	6.65E-02	2.47E-01	-1.5	18.1	53.0	16.8	45.6
PPP4R4.b	5.76E-04	1.85E-01	-1.5	6.1	17.6	6.5	18.0
MTMR4	1.83E-02	2.25E-01	-1.5	5.6	15.0	5.3	16.0
NDE1	3.90E-03	2.08E-01	-1.5	8.6	23.4	8.1	24.2
MT-ND1	3.40E-02	2.32E-01	-1.6	65.3	205.5	86.7	242.0
ZBTB32	1.74E-01	2.87E-01	-1.6	1.4	11.0	6.6	12.8
IL6ST	2.48E-02	2.25E-01	-1.6	3.2	11.6	5.1	13.3
NDE1.b	6.56E-02	2.45E-01	-1.6	8.6	23.3	8.1	27.0
MAP3K5	1.10E-01	2.59E-01	-1.6	5.0	22.3	7.8	16.1
CAPG	7.08E-04	1.85E-01	-1.6	22.6	71.8	24.3	70.0
SNX11	1.47E-02	2.25E-01	-1.6	7.1	23.7	9.6	26.7
CCND1	1.83E-02	2.25E-01	-1.6	4.6	14.2	4.6	13.7
CD40	1.79E-03	1.88E-01	-1.6	52.5	176.3	59.6	169.5
TRIM22	3.58E-02	2.33E-01	-1.6	24.3	88.0	32.0	86.6
IRF9	4.80E-02	2.37E-01	-1.7	9.1	31.0	9.6	27.8
IFIT3	8.16E-02	2.50E-01	-1.7	5.8	17.9	6.8	22.0
SPIC	1.77E-02	2.25E-01	-1.7	2.7	10.2	3.8	10.8
HLA-A.g	3.16E-02	2.31E-01	-1.7	13.3	47.0	18.3	55.3
ARID5A	4.71E-02	2.37E-01	-1.7	2.9	12.6	4.2	10.2
PPP1R1C	2.87E-02	2.28E-01	-1.7	16.3	45.5	13.6	51.5
MVP	2.35E-03	1.97E-01	-1.7	4.3	15.5	4.9	14.8
NCF4	5.86E-04	1.85E-01	-1.7	7.3	24.1	7.2	24.3

SAMSN1	1.01E-02	2.19E-01	-1.8	36.7	131.1	39.9	130.7
QTRT1	1.30E-01	2.66E-01	-1.8	5.9	13.4	3.7	19.5
RASSF4	7.80E-04	1.85E-01	-1.8	14.7	52.5	15.6	53.0
B4GALT5	1.25E-03	1.87E-01	-1.8	3.0	10.9	3.4	11.3
DUSP10	7.60E-03	2.08E-01	-1.8	3.5	15.1	4.6	13.8
IRF1	3.61E-03	2.08E-01	-1.8	11.4	40.8	11.0	39.9
IL4I1	2.00E-02	2.25E-01	-1.9	3.0	12.0	3.7	12.2
B4GALT6	2.49E-01	3.19E-01	-1.9	4.8	18.8	3.0	9.8
STAT1	5.60E-04	1.85E-01	-1.9	19.5	75.3	20.9	73.6
SLAMF1	2.24E-02	2.25E-01	-1.9	9.5	36.3	9.7	34.4
SLAMF7	2.90E-03	2.07E-01	-1.9	2.6	10.3	2.9	9.9
ZNF33B	7.59E-02	2.49E-01	-1.9	8.2	37.8	10.2	30.4
TXLNB	3.07E-02	2.29E-01	-1.9	12.4	41.4	9.5	41.4
HLA-E.f	2.38E-01	3.13E-01	-1.9	6.8	21.5	1.6	10.6
ADAMDEC1	3.23E-02	2.32E-01	-1.9	2.5	11.4	3.4	11.5
BCL2A1	8.53E-03	2.10E-01	-2.0	63.1	245.0	65.9	254.1
ICAM1	4.59E-02	2.37E-01	-2.0	14.0	58.3	14.5	52.2
ATF3	1.09E-03	1.87E-01	-2.0	3.3	13.0	3.2	13.0
LMNA	6.88E-02	2.47E-01	-2.0	8.4	36.8	8.5	31.2
FABP6	6.19E-02	2.42E-01	-2.0	10.1	43.6	9.8	37.6
IL7R	2.63E-02	2.26E-01	-2.1	1.9	11.2	4.0	13.5
RASGEF1B	1.26E-02	2.23E-01	-2.1	7.2	35.0	9.5	35.9
PDE4DIP	9.60E-02	2.55E-01	-2.1	6.3	25.8	7.5	33.2
TMEM120A	1.55E-01	2.77E-01	-2.1	5.5	17.7	4.7	26.2
GOPC	1.25E-03	1.87E-01	-2.1	10.5	49.7	12.1	48.6
SAMD9L	6.08E-03	2.08E-01	-2.1	2.6	13.1	3.6	13.8
GPR137B	9.45E-02	2.55E-01	-2.1	5.1	19.7	5.1	24.9
TRAF1	2.77E-02	2.28E-01	-2.1	8.5	38.1	9.9	42.4
BCL2	1.08E-02	2.21E-01	-2.2	4.5	20.0	4.3	19.2
AL132671.2	2.13E-02	2.25E-01	-2.2	9.5	43.6	9.8	46.2
TDRD7	2.90E-03	2.07E-01	-2.2	5.0	25.2	5.9	25.6
EPSTI1	3.37E-03	2.08E-01	-2.3	54.3	259.8	51.6	259.4
FLNB	3.76E-01	3.72E-01	-2.3	2.6	19.9	2.7	6.0
BCL2L1	1.74E-03	1.88E-01	-2.3	55.5	292.8	61.6	289.9
CLIC2	2.96E-03	2.07E-01	-2.3	6.0	34.9	8.3	36.9
NFKBIZ	1.28E-02	2.23E-01	-2.4	5.5	31.9	7.3	35.2
GADD45B	6.17E-02	2.41E-01	-2.4	2.8	15.4	2.7	13.2
NFKBIA	4.21E-03	2.08E-01	-2.4	31.0	168.3	34.3	175.1
MGLL	4.15E-02	2.36E-01	-2.4	2.2	10.1	1.8	11.5
SLFN5	6.81E-04	1.85E-01	-2.5	2.5	14.3	2.8	14.8
NUS1	4.45E-02	2.37E-01	-2.5	4.4	25.6	5.1	29.2
FCRL4	2.51E-03	2.03E-01	-2.6	1.9	12.7	2.5	13.2
CRIP1	5.03E-02	2.37E-01	-2.6	24.8	157.8	33.9	188.6
GBP1	4.14E-02	2.36E-01	-2.6	3.6	24.3	5.0	28.6
NCF2	1.45E-03	1.87E-01	-2.7	8.2	59.1	11.0	61.7

TFEC	4.16E-03	2.08E-01	-2.7	17.5	126.9	22.8	135.2
IFI44L	7.36E-02	2.48E-01	-2.7	9.6	58.9	10.4	72.0
ZBTB10	5.56E-02	2.38E-01	-2.7	5.3	41.7	6.4	35.4
AL928654.4	1.92E-03	1.92E-01	-2.7	143.5	1108.0	198.4	1148.9
ZNF782	4.91E-01	3.73E-01	-2.7	1.7	24.2	2.3	2.3
INSIG1	2.81E-02	2.28E-01	-2.8	26.7	178.6	23.1	162.3
ACSBG1	2.02E-02	2.25E-01	-2.8	1.7	17.3	2.8	15.1
KDM2B	1.23E-02	2.23E-01	-2.9	9.1	65.2	9.0	67.4
CLDN12	4.41E-01	3.73E-01	-2.9	2.0	29.2	2.4	4.7
LGALS1	1.96E-03	1.92E-01	-3.1	12.1	133.7	19.3	127.7
TRIP10	1.07E-01	2.58E-01	-3.1	9.8	84.7	7.1	61.8
PTK2B	2.64E-03	2.04E-01	-3.3	7.4	79.4	8.6	77.0
CD83	3.38E-02	2.32E-01	-3.5	21.6	275.7	26.1	247.3
ESR1	4.37E-01	3.73E-01	-3.5	0.6	2.7	1.3	18.3
TNFAIP3	8.44E-02	2.52E-01	-3.6	11.4	186.1	16.2	144.1
MYEOV	6.74E-03	2.08E-01	-3.6	0.9	11.7	1.1	12.3
CCR7	1.35E-02	2.23E-01	-3.7	3.7	47.9	4.1	50.3
IL32	1.70E-02	2.25E-01	-3.7	1.7	21.0	1.5	19.7
LNX1	2.34E-02	2.25E-01	-3.8	1.2	15.7	1.1	14.6
LAMP3	2.58E-02	2.25E-01	-3.9	0.8	14.2	1.1	12.9
CCL22	2.62E-02	2.26E-01	-3.9	1.0	18.1	1.6	20.2
MTSS1	2.10E-02	2.25E-01	-4.0	1.4	27.8	2.2	30.6
TTC39A	1.70E-02	2.25E-01	-4.0	1.2	21.2	1.4	20.0
NRG4	1.78E-03	1.88E-01	-4.0	3.5	60.8	3.8	60.8
EBI3	3.52E-03	2.08E-01	-4.1	2.9	58.2	4.0	57.9
TPCN2	3.75E-02	2.33E-01	-4.2	0.7	11.0	0.5	9.8
SRGN	1.23E-03	1.87E-01	-4.5	33.2	816.3	37.7	804.9
UBD.e	5.42E-02	2.37E-01	-4.8	2.7	72.7	2.3	61.6
UBD.b	7.10E-02	2.48E-01	-5.1	3.6	186.8	6.6	149.4
IFI27	2.81E-02	2.28E-01	-5.1	1.2	41.8	1.2	38.3
CXCL10	1.20E-02	2.23E-01	-5.2	2.9	118.9	3.7	124.1
TNFRSF9	4.18E-02	2.36E-01	-5.3	0.3	14.3	0.5	16.3
IFI27.b	3.22E-02	2.32E-01	-5.4	1.2	49.3	1.2	54.5
GRIA4	1.91E-02	2.25E-01	-5.6	0.5	30.3	0.8	32.4
AP003071.5	1.29E-01	2.65E-01	-6.2	0.2	18.2	0.2	12.1
TNFAIP6	3.48E-02	2.33E-01	-7.2	0.1	10.6	0.1	9.5
SERPINB10	9.20E-03	2.18E-01	-7.3	0.1	12.1	0.1	11.7
ZC3H12C	2.50E-02	2.25E-01	-7.6	0.1	11.9	0.1	12.9

Table 15: Upregulated genes in the presence of LMP1-induction found via RNA Seq

Mass Spectrometry Data

Table Mass Spectrometry No LMP1 Group

Gene	Unique Peptides	Sequence Coverage	Control MS Count	No LMP1 MS Count N=1	No LMP1 MS Count N=2
KIFC1	36	55.6	0	57	44
MORC2	27	29.6	0	45	41
TRIP12	30	20.3	0	45	26
RTCB	17	39.2	0	25	38
OAS3	25	25	0	38	24
MAP4	26	16.1	0	29	27
WIZ	18	21.5	0	21	23
C14orf166	8	39.3	0	22	17
HCFC1	13	8.6	0	18	19
ATP5B	14	39.5	0	20	17
RCC1	15	54.6	0	24	12
SKIV2L2	17	18.2	0	18	15
CDC27	1	24.8	0	18	14
EBNA1BP2	9	31.9	0	21	11
MED23	16	12.7	0	18	12
FAM98B	6	25.6	0	14	14
ABCF1	13	19.7	0	14	13
KIF2C	12	21.8	0	13	14
MED24	11	17.7	0	16	8
ZC3H11A	12	18.6	0	12	12
DDX50	13	27.7	0	14	10
HMMR	10	22	0	8	14
RPP30	8	35.8	0	12	10
EXOSC2	8	36.9	0	11	11
SYMPK	13	14.5	0	10	11
FEN1	8	29.7	0	10	11
TRIM25	13	28.1	0	8	13
CENPV	7	37.1	0	9	12
DTX3L	11	21.9	0	13	7
C2orf49	7	41.8	0	8	12
CMAS	9	27.6	0	13	6
SNIP1	7	22.7	0	10	8
CDKN2AIP	9	23.1	0	9	9
EXOSC9	7	22.7	0	8	9
MED14	9	9.7	0	8	9
ERCC2	11	20	0	12	4
EXOSC6	5	29	0	6	10
MORF4L1	8	25.6	0	4	11
DRG1	8	29.2	0	8	7

EXOSC4	5	26.5	0	8	7
WDR33	7	9.9	0	7	7
CCNK	4	11.2	0	9	5
CPSF3	6	11.1	0	5	9
EXOSC7	7	34.7	0	10	4
CGGBP1	5	34.1	0	7	6
PSPC1	6	22.9	0	3	10
EXOSC5	4	19.1	0	6	7
EXOSC3	5	28.7	0	8	4
RRBP1	12	10.9	0	3	9
MED12	6	3.9	0	7	4
ACTL6A	6	22.1	0	10	1
KPNB1	5	8	0	6	5
CDC23	8	21.3	0	9	2
TPX2;HCA90	6	10.2	0	4	7
PARP9	7	13.9	0	5	5
MED16	6	9.4	0	7	3
GTF2F1	5	13.9	0	2	8
MED4	6	27.8	0	6	4
NCAPD3	8	7	0	9	1
AKAP17A	7	11.9	0	9	1
ARHGEF2	6	8	0	4	6
SCHIP1;IQCJ-SCHIP1;IQCJ	2	5.6	0	5	4
PHB2	4	19.9	0	7	2
GTF2H2C;GTF2H2	5	13.9	0	6	3
CDC16	6	17.5	0	5	4
GTF2H4	5	15.8	0	6	3
HDGFRP2	4	9.7	0	4	5
EED	5	14.8	0	6	2
RPL15	6	32.2	0	7	1
LSM14A	4	12.8	0	4	4
SERBP1	3	10.9	0	5	3
FAM98A	5	20.8	0	5	3
HEATR1	5	2.7	0	5	2
PXK	4	8.1	0	3	4
CPSF4	3	25.1	0	5	2
CWC22	5	6.5	0	6	1
MLLT1	5	11.4	0	1	6
FRG1	2	25	0	5	2
ANAPC5	4	7.7	0	3	4
EBF1	3	9.4	0	6	1
MED17	3	4.6	0	3	4
OSBPL1A	3	51.2	0	4	2
RTEL1	1	1.4	0	4	2

MAGOH;MAGOHB	1	54.8	0	5	1
MED22	2	22.5	0	3	3
TADA2B	3	9.5	0	5	1
PPP1R18	2	2.4	0	2	4
CDCA5	3	9.8	0	2	4
RNF40	5	9.9	0	5	1
MED15	1	4.1	0	1	5
ZC3H15	4	13.4	0	2	4
BOD1L1	5	2.2	0	2	4
MPHOSPH6	3	23.9	0	3	3
PHF5A	4	40.9	0	5	1
PHB	3	17.5	0	5	1
RPP38	5	18.7	0	3	3
CBX3	4	30.1	0	4	2
CDK13	3	5	0	3	3
NELFB	2	4.3	0	3	2
TBL3	1	7.2	0	4	1
MED21	2	20.1	0	3	2
HM13	2	6.6	0	4	1
SLC25A3	3	10.5	0	3	2
RBMX2	3	10.2	0	3	2
DHX37	3	3.8	0	2	3
RPP40	2	23.9	0	4	1
RSRC1	4	21	0	4	1
CCHCR1	2	4.4	0	1	4
PRDM2	3	2.7	0	3	2
SLC25A4	3	28.9	0	3	2
CACNA1A	2	4.1	0	4	1
ANAPC1	4	2.5	0	2	3
CPSF2	4	7.5	0	2	3
CDK19;CDK8	2	10	0	2	2
WDR3	1	1.1	0	2	2
KLF12	2	7	0	2	2
PBRM1	3	2.3	0	3	1
CD2BP2	2	6.5	0	3	1
MED20	3	17	0	2	2
ALKBH5	3	12.7	0	1	3
INTS3	3	6.4	0	3	1
MED6	3	16.9	0	2	2
NIPBL	3	1.2	0	2	2
EP400	4	1.7	0	3	1
RBMXL1	3	28.5	0	1	3
NFIA;NFIX	2	7.1	0	3	1
GTF2F2	2	9.6	0	1	3
RFX5	2	11.5	0	3	1

RRP12	1	1.3	0	3	1
CDC27	1	26.9	0	1	2
TRMT6	2	9.3	0	1	2
NELFA;WHSC2	1	6.7	0	2	1
ENY2	2	27.1	0	2	1
ANAPC4	2	4.9	0	2	1
PUF60	0	49.6	0	2	1
CBX4	1	3.2	0	2	1
FAM76A	3	12.5	0	2	1
MNAT1	1	6	0	1	2
CDC26	1	14.1	0	1	2
WHSC1L1	2	3.9	0	2	1
ANAPC10	1	11	0	1	2
EXOSC1	3	20.5	0	1	2
GTF2E2	1	7.1	0	2	1
ANAPC2	2	4	0	2	1
TRRAP	38	12.4	0	48	9
EXOSC10	22	35.2	0	32	24
DNMT1	51	38.5	0	93	60
DDX49	14	34.6	0	24	20
MED1	19	20.3	0	21	18
DDX1	20	34.3	0	31	34
NDNL2	13	43.1	0	24	9
PRC1	14	27.9	0	23	13
CPSF1	17	15	0	11	13
NSMCE2	6	32.6	0	11	12
XRCC1;FOXH1	7	15.2	0	10	8
MINA	7	19.8	0	10	5
NSMCE1	6	27.4	0	10	3
GTPBP1	10	17.9	0	9	11
AFF4	7	7.7	0	9	9
ATRIP	12	20.7	0	9	6
DDB1	8	9.6	0	9	3
GAPDH	5	26.3	0	9	2
NSMCE4A	7	24.2	0	14	6
CDK12	3	4.8	0	7	5
MMTAG2	4	25.9	0	7	2
SMC6	29	35.7	0	44	20
KNOP1	9	26.6	0	11	5
MDC1	4	3.6	0	5	1
VRK1	16	47.7	0	34	22
NOP14	3	3.2	0	4	3
BUD31	5	32.2	0	7	5
C10orf12	3	3.7	0	7	3
PCBP1	4	21.6	0	10	10

SETX	7	3.4	0	3	5
CENPU;MLF1IP	4	11.7	0	3	5
UPF3B	4	14	0	3	4
APRG1	1	5.3	0	5	7
RSBN1	7	13.8	0	5	5
KTN1	3	2	0	5	3
ATR	66	34.9	0	86	75
SMC5	31	29.1	0	43	29
DDX18	21	40.6	0	43	22
HNRPLL;HNRNPLL	6	34.4	0	15	17
RFC1	24	27.5	0	44	29
RBM26	12	18.5	0	13	13
RSBN1L	10	15.5	0	12	7
SMCHD1	81	46	0	199	100
TCOF1	0	20.5	0	49	35
YBX1;YBX3	11	43.8	0	18	22
ATAD5	6	4.5	0	4	3
NCAPG2	12	13.8	0	12	3
RIF1	6	4.2	0	4	3
APEX1	16	67.3	0	37	18
NOP2	18	25.7	0	22	17
DIDO1	23	13.4	1	31	20
THOC6	13	59.8	1	19	15
ATP5A1	11	31.4	1	18	6
ERCC3	11	21	1	11	9
VDAC2	8	32.5	1	14	8
THOC7	7	42.2	1	8	11
SRP68	7	18.5	1	4	8
POLDIP3;PDIP46	7	23.3	1	10	2
THOC3	8	35.3	1	5	6
SNRPG;SNRPGP15	2	20.3	1	6	2
CTR9	5	6.1	1	3	5
CANX	4	8.4	1	6	2
NKAP	3	14.5	1	3	4
SNRPF	3	39.5	1	3	2
FAM133B;FAM133A	3	35.2	1	3	2
ACSL1	12	20.8	1	14	8
H2AFV;H2AFZ	3	31.2	1	6	3
PSIP1	15	29.4	1	35	26
RBM15	14	21.1	1	14	12
ZNF638	16	11.7	1	19	11
SFSWAP	7	11.5	1	5	7
IGF2BP1	6	18.5	1	15	7
POP1	18	24.4	1	29	25
IK	11	28.5	1	13	3

CCAR1	25	30	1	48	14
THOC2	39	31	2	74	47
THOC5	22	34	2	33	26
THOC1	21	36.1	2	34	24
TMPO	22	42.2	2	31	23
HSPA5	14	28.7	2	17	12
SLC25A5	5	32.6	2	15	12
NOLC1	7	12	2	8	12
HELLS	10	14.7	2	14	6
SNRNP27	4	29.7	2	9	8
MFAP1	4	15.3	2	8	4
DYNLL1;DYNLL2	3	49.4	2	7	5
MKI67	116	46	2	206	153
THRAP3	12	17.8	2	20	24
SRP72	12	23.7	2	12	10
RAD21	10	31.1	2	19	5
DKC1	12	32.9	2	14	11
PDS5B	29	28.2	2	51	21
SNRPA	3	20.2	3	5	4
RRP1B	24	40.3	3	34	28
HIST1H1C	3	39.9	3	43	20
SRSF9	7	33.5	3	16	7
MYEF2	9	16.8	3	7	5
SAFB	13	30.3	3	28	18
RBBP6	13	9.2	3	8	11
VDAC1	10	43.5	4	12	12
DDX3X;DDX3Y	9	19	4	12	9
SF3B5	3	43	4	8	7
SNRPE	2	25	4	6	7
BRD2	14	29.7	4	28	17
PRPF19	14	39.5	5	27	9
SNRNP40	13	48.7	5	24	10
FIP1L1	10	26.2	5	10	11
SREK1	12	27.1	5	26	14
BRD3	8	28.2	5	12	12
SLTM	16	19.1	5	31	30
SNRPD1	3	37.8	5	7	10
SMU1	14	33.5	5	19	12
CDC5L	23	41.1	5	33	11
NFKB1	6	10.5	6	7	7
RUVBL2	17	39.1	6	22	13
SYNCRIP	21	46.2	6	48	53
DDX41	19	35.3	7	24	14
BRD4	12	15.6	7	33	30
DDX17	17	39.5	7	48	19

BUB3	12	48.6	7	15	16
RCC2	28	64.2	7	83	70
PUF60	1	52.8	8	44	34
BCLAF1	13	23.8	8	18	16
ZFR	19	23.5	8	20	14
PRPF38B	15	32.8	8	28	23
NPM1	12	49.8	8	15	13
PDS5A	43	42.3	9	93	35
PPIG	11	14.5	9	19	15
XRN2	11	15.2	10	28	12
SRSF10	12	58.2	10	24	12
U2AF2	12	42.9	11	36	29
HNRNPA1;HNRNPA1L2	18	54.1	12	43	66
ACIN1	25	24.9	15	68	21
NUMA1	42	27.2	16	63	34
ILF2	17	51.3	16	45	26
LUC7L2	0	32.1	16	29	17
HNRNPA2B1	23	66.3	16	66	90
SRRM1	1	23.4	16	66	35
DDX21	38	51.5	20	106	60
IFI16	22	53.9	21	102	97
HNRNPU	45	50.1	29	170	149
HNRNPC;HNRNPCL1	8	40.8	30	63	59
STRBP	26	47.7	30	46	41
DDX23	30	38.8	35	55	41
DDX46	42	43.5	38	60	54
HNRNPR	2	48	42	109	107
SRRM2	68	30.7	44	153	58
EFTUD2	48	62.4	48	108	57
SF3B1	43	39.2	49	69	63
DDX5	26	51.1	52	106	93
PRPF8	96	47.7	85	206	107

Table 16: The 293 specific proteins found in the group No LMP1 from Mass Spectrometry pull-down during DNA Affinity chromatography

Table Mass Spectrometry LMP1 Group

Gene	**Unique Peptides**	**Sequence Coverage**	**Control MS Count**	**LMP1 MS Count N=1**	**LMP1 MS Count N=2**
SUPT6H	25	18.8	0	35	5
TBL2	13	43.1	0	22	1
RBM12B	6	7	0	11	11
ZNF106	12	17.9	0	18	4

USF2	4	26.1	0	13	8
CHD1L	11	18.7	0	11	8
TXLNA	10	22	0	17	2
NACA	3	59.2	0	12	6
NOL6	11	12.9	0	16	2
SDAD1	10	19.2	0	14	4
IRF2	5	50.4	0	9	8
URB1	13	8.1	0	14	2
RPL10	8	39.7	0	8	8
LLPH	2	13.2	0	7	8
AKAP8L	3	4.6	0	9	5
GTF2B	7	26.9	0	12	1
PURB	5	18.3	0	5	8
RARA;RARB	2	5.4	0	6	6
MTA2	8	19.9	0	11	1
MDN1	4	0.8	0	7	4
RPL29	3	20.8	0	7	4
DDX51	6	11.7	0	7	2
ZBTB7B	4	6.5	0	6	3
PRRC2B	2	1.7	0	6	2
BTF3	3	29.6	0	7	1
COPA	4	4	0	2	6
LOC102288414	3	23.2	0	5	3
ZBTB48	3	5.7	0	6	2
FER1L4	1	1	0	5	2
THAP1	4	21.1	0	5	2
CRAMP1L	4	4.4	0	6	1
CHAMP1	4	4.8	0	5	2
NOP16	3	24.2	0	4	3
ZNF800	3	4.5	0	5	2
MYNN	3	14.7	0	3	4
ZHX1	3	3.8	0	4	2
EBI3	2	8.3	0	5	1
PARP14	6	5.5	0	5	1
LARP4B	2	3.1	0	4	2
ABI1	2	5.8	0	4	2
ELF1	1	5	0	3	2
PHAX	3	8.1	0	4	1
ZNF770	3	30.3	0	3	2
PRDM15;ZNF298	3	2.7	0	3	2
RRP7A	2	8.6	0	4	1
MYSM1	3	3.6	0	4	1
NCAPH	1	2	0	3	2
RREB1	2	1.1	0	2	3
ZCRB1	3	24.4	0	4	1

TNRC18	3	1.2	0	3	2
MAK16	3	11.7	0	4	1
CCDC106	2	8.6	0	2	3
ZBTB7A	2	6.7	0	1	4
SUPT4H1	2	30.3	0	3	1
KIAA1429	1	1	0	3	1
COPB2	4	7.3	0	1	3
ZNF668	1	2.1	0	2	2
TOP3A	2	6	0	2	2
RPL34	2	12.8	0	3	1
ZNF91;ZNF595	2	2.6	0	3	1
CPB1	1	3.4	0	2	2
CWF19L2	2	3.2	0	1	3
BCL11A	3	5.4	0	2	2
ZNF93	3	7.5	0	3	1
SHOC2	2	4.3	0	3	1
ZNF286A;ZNF286B	2	5.1	0	2	2
ZFP62	2	3	0	3	1
ZNF24	1	3.5	0	2	2
RPL26L1	1	44.1	0	2	1
CHRAC1	2	46.4	0	2	1
ZBTB17	3	16.6	0	2	1
ZNF687	2	1.4	0	1	2
ZNF593	3	40.3	0	1	2
REST	2	2.3	0	1	2
NOL12	2	10.3	0	2	1
HIRIP3	5	17.3	0	21	8
RPL8	5	19.5	0	22	8
STK10	11	15.5	0	10	4
ABT1	6	20.6	0	10	4
FAM60A	4	33.9	0	10	3
RPL32	6	42.1	0	9	3
RBM19	6	6.4	0	17	6
NOC3L	4	7.6	0	7	1
ZNF768	6	13.2	0	13	10
NEIL1	6	22.8	0	6	3
RRP15	2	6.7	0	6	3
RPL13	4	31.7	0	6	2
VRK3	3	7.1	0	5	2
HBS1L	14	39.8	0	20	1
USP42	2	2.5	0	4	1
HOXB4	6	18.3	0	12	12
ERI1	5	21.2	0	6	2
ZNF652	4	6.6	0	3	2
TRIM4	1	5.4	0	3	1

SMNDC1	2	13.9	0	3	1
	2	17.8	0	7	4
NOC2L	19	23.9	0	45	30
TREX1	12	39.1	0	20	17
RPL27A	8	39.9	0	27	34
RPL23	10	72.9	0	41	15
RPL36A;RPL36AL;RPL36A-HNRNPH2	6	29.6	0	12	13
PPAN	17	38.5	0	44	12
CEBPZ	28	29.1	0	57	32
CTCF	17	21	0	45	37
BANF1	11	80.9	0	31	21
RPL35	6	35.8	0	20	27
YY1;YY2	10	21	0	29	20
RPL17	9	47.4	0	25	13
WDR76	19	32.3	0	33	19
UHRF2	9	14	0	12	6
CPSF3L	2	4.4	0	6	6
NCKAP1L	10	10.8	0	17	6
RPL26;KRBA2;RPL26L1	1	60.7	0	25	53
PDCD11	72	39.3	0	225	168
WBP11	2	15.5	0	10	11
DDX27	15	18.7	0	30	12
HOXC4	5	22.3	0	8	4
MYBBP1A	13	12	0	13	4
SUPT5H	19	23.5	0	28	15
ZNF629	10	13.9	0	19	11
REL	13	26.6	0	24	14
RPL22	5	62.1	0	4	7
RPL23A	14	44.3	0	55	61
FTSJ3	13	17.5	0	34	9
RPF2	4	17.6	0	5	3
RBM28	26	31.5	0	42	21
ISG20L2	5	19.3	0	11	5
RPL7L1	11	37.4	0	20	10
CMSS1	17	53.3	0	45	26
SP110	4	9.9	1	4	3
CHD1	9	5.5	1	13	10
MAP4K4;TNIK;MINK1	7	9.2	1	7	3
CDK11A;CDK11B	5	5.4	1	6	3
HJURP	5	9.6	2	3	3
RPS9	10	37.6	2	12	7
RPL39P5;RPL39	3	23.5	2	7	7
RPL38	7	57.1	3	32	23
HSPB1	6	31.7	4	22	14

RPS26;RPS26P11	3	31.3	4	13	6
TOP1	53	53.6	4	217	225
RPS23	3	16.1	5	11	11
TOP2B	34	46.4	7	57	40
RPS6	14	41	9	45	30
RPS4X	16	55.5	11	40	16
UBTF	45	47	14	133	156

Table 17: The 144 specific proteins found in the group LMP1 from Mass Spectrometry pull-down during DNA Affinity chromatography